Special Issue

# Transplant International

## Clinical trial design and endpoints in kidney transplantation

Transplant International

frontiers | Publishing Partnerships

# Clinical Trial Design and Endpoints in Kidney Transplantation

In 2018, the European Society for Organ Transplantation (ESOT) initiated a Broad Scientific Advice request to the European Medicines Agency (EMA), to discuss their Committee for Medicinal Products for Human Use (CHMP) guidelines relating to the development of innovative therapies in the field of kidney transplantation.

The request suggested several options for redefining endpoints used in clinical studies of kidney transplantation. It also proposed study design elements for stratification and subsequent employment of these novel proposals into future clinical trials. The proposed endpoints also aimed to be suitable for the conditional marketing authorization pathway. The EMA subsequently evaluated the discussion document and evidence provided by ESOT, issuing their final recommendations in December 2020.

This Special Issue of Transplant International provides a detailed summary of key aspects of the proposals and recommendations for definitions of endpoints for clinical trials of kidney transplantation.



ISSN 1432-2277
ISBN 978-2-88976-352-8
DOI 10.3389/978-2-88976-352-8

Transplant International

# Table of contents

ESOT | Transplant International

EXECUTIVE SUMMARY
published: 20 May 2022
doi: 10.3389/ti.2021.10142

# Redefining Risk Stratification and Endpoints for Clinical Trials in Kidney Transplantation: Rationale and Methodology of Proposals Submitted to the European Medicines Agency by the European Society for Organ Transplantation

*Maarten Naesens*[1]*, *Stefan Schneeberger*[2] *and the ESOT Working Group*

[1]*Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium,* [2]*Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria*

The European Society for Organ Transplantation (ESOT) submitted a Broad Scientific Advice request to the European Medicines Agency (EMA) in 2018, to explore whether updating guidelines on clinical trial endpoints would encourage innovations in kidney transplantation research, thereby improving long-term outcomes for allograft recipients. The request was refined collaboratively by the EMA and ESOT, with the EMA issuing a final response in December 2020. This *Transplant International* special issue explores the topics that were the focus of these interactions between the EMA and ESOT. Articles explore the current issues and dilemmas in kidney transplantation, primarily relating to unclear or outdated risk stratification and markers of transplantation success, although several potential improvements for outcomes assessment are also suggested. Discussions between the EMA and ESOT and recommendations are summarized, in the hope that this project will generate further discussion eventually generating a consensus on clinical trial endpoints and risk stratification, increase the quality of research in transplantation medicine, and improve long-term outcomes for kidney transplant recipients.

**Keywords: kidney transplantation outcome, EMA guideline, efficacy endpoint, long-term outcome, improvement, European Society for Organ Transplantation, risk stratification**



***Correspondence:***
*Maarten Naesens*
*maarten.naesens@kuleuven.be*

***Received:*** *21 October 2021*
***Accepted:*** *15 November 2021*
***Published:*** *20 May 2022*

***Citation:***
*Naesens M, Schneeberger S and the ESOT Working Group (2022) Redefining Risk Stratification and Endpoints for Clinical Trials in Kidney Transplantation: Rationale and Methodology of Proposals Submitted to the European Medicines Agency by the European Society for Organ Transplantation. Transpl Int 35:10142. doi: 10.3389/ti.2021.10142*

## INTRODUCTION

Over many decades, progress in the treatment of acute rejection markedly improved the short-term success of kidney allografts, such that graft survival in the first year after transplantation now exceeds 90% [1]. However, improvement rates have decelerated. Data from 135 kidney transplant centers in 21 European countries (187,787 individual transplantations) indicate that the improvement of graft survival has slowed significantly since 2000, even when considering the increased age of donors and recipients [1]. Initiatives to further improve graft survival rates are therefore needed, but the nature of such initiatives is an important point for discussion.

Certainly, several pharmaceutical regimens have been developed, such as efficacious and relatively well-tolerated immunosuppressants, which have enabled very good short-term outcomes to be achieved in patients (and with an acceptable risk of graft rejection). However, the ensuing misconception is that all major hurdles in transplantation have been overcome [2, 3]. Good short-term outcomes that are observed in transplantation recipients do not always translate into satisfactory long-term graft functioning or patient-survival rates [4, 5]. Lack of long-term success creates difficulty in defining suitable surrogate endpoints for clinical trials, which is problematic for ongoing research and could discourage academic/commercial investment in kidney transplantation [3].

Consequently, while clinical progress in kidney transplantation is slowing down, rates of allograft loss continue to be unacceptably high. The extensive negative impact that this has on patient health and well-being—as well as the high long-term health-associated cost burden [6]—clearly indicates a need for novel, effective management strategies, tested according to endpoints that suit current practice and regulations. In addition, there are no approved surrogate markers for long-term graft failure in kidney transplantation, necessitating long-term interventional studies as an urgent priority.

Innovations in kidney transplantation often focus on the prevention and treatment of acute allograft rejection [7]. While current immunosuppressive regimens have reduced the incidence of rejection in low-risk organ recipients [8], many patients have a high immunological risk and could benefit from better preventive and therapeutic options than those available [9]. Stratification of patients and allografts according to immunological risk, however, is not standardized.

## CURRENT EMA GUIDANCE ON CLINICAL STUDIES FOR KIDNEY TRANSPLANTATION

Released in 2008, the European Medicines Agency (EMA) guideline (CHMP/EWP/263148/06) [10] provides guidance on the conduct of clinical studies for solid organ transplantation (not specific for kidney transplantation) by defining treatment goals, study designs, outcome measures, and data analyses for new immunosuppressive therapies developed to prevent and treat allograft rejection. The guideline [10] defines the primary efficacy endpoint for novel immunosuppressants in solid organ

**BOX 1 | Individuals involved with the EMA-CHMP request for Broad Scientific Advice project, on behalf of the European Society for Organ Transplantation.**

**COORDINATOR**

Maarten Naesens: Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

**CONTRIBUTORS**

Jan Ulrich Becker: Institute of Pathology, University Hospital Cologne, Cologne, Germany
Maria Irene Bellini: Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy
Oriol Bestard: Department of Nephrology and Kidney Transplantation, Vall d'Hebrón University Hospital, Barcelona, Spain
Georg A. Böhmig: Division of Nephrology and Dialysis, Department of Internal Medicine, Medical University of Vienna, Vienna, Austria
Klemens Budde: Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany
Fergus Caskey: Department of Population Health Sciences, University of Bristol, Bristol, United Kingdom
Frans Claas: Eurotransplant Reference Laboratory, Department of Immunology, Leiden University Medical Center, Leiden, Netherlands
Lionel Couzi: Department of Nephrology, Transplantation and Dialysis, Bordeaux University Hospital, Bordeaux, France
Fritz Diekmann: Department of Nephrology and Kidney Transplantation, Hospital Clinic Barcelona, Barcelona, Spain
Fabienne Dobbels: Academic Center for Nursing and Midwifery, Department of Public Health and Primary Care, KU Leuven, Leuven, Belgium
Lucrezia Furian: Kidney and Pancreas Transplantation Unit, University of Padua, Padua, Italy
Denis Glotz: Paris Translational Research Center for Organ Transplantation, Hôpital Saint Louis, Paris, France
Josep Grinyó: Emeritus Professor of Medicine, University of Barcelona, Barcelona, Spain
Uwe Heemann: Department of Nephrology, Technical University of Munich, Munich, Germany
Dennis Hesselink: Erasmus MC Transplant Institute, University Medical Center Rotterdam, Rotterdam, Netherlands
Luuk Hilbrands: Department of Nephrology, Radboud University Medical Center, Nijmegen, Netherlands
Ina Jochmans: Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium
Alexandre Loupy: Paris Translational Research Center for Organ Transplantation, Hôpital Necker, Paris, France
Nizam Mamode: Department of Transplantation, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom
Rainer Oberbauer: Department of Nephrology and Dialysis, Medical University of Vienna, Vienna, Austria
Liset Pengel: Centre for Evidence in Transplantation, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom
Marion Rabant: Department of Pathology, Hôpital Necker-Enfants Malades, Paris, France
Marlies Reinders: Erasmus MC Transplant Institute, Department of Internal Medicine, University Medical Center Rotterdam, Rotterdam, Netherlands
Lionel Rostaing: Department of Nephrology, Dialysis and Transplantation, Toulouse University Hospital, Toulouse, France
Candice Roufosse: Centre for Inflammatory Disease, Department of Immunology and Inflammation, Imperial College London, London, United Kingdom
Stefan Schneeberger: Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria
Daniel Seron: Department of Nephrology and Kidney Transplantation, Vall d'Hebrón University Hospital, Barcelona, Spain
Olivier Thaunat: Department of Transplantation, Nephrology, and Clinical Immunology, Edouard Herriot Hospital, Hospices Civils de Lyon, Lyon, France
Allison Tong: Sydney School of Public Health, The University of Sydney, Sydney, Australia

**BOX 2 | Questions presented to EMA by ESOT in their request for Broad Scientific Advice on clinical trial design and endpoints in kidney transplantation. This special issue presents current knowledge and perspectives from the European Society for Organ Transplantation (ESOT) on these five core questions, based on evidence and clinical and research experience.**

Q1: Does CHMP agree with the updated definitions of rejection and their potential use as primary endpoints in studies of kidney transplantation?

Q2: Does the CHMP agree with the proposed definitions of allograft (dys)function in kidney transplantation, and the recommendations for parameters that could be used as primary endpoints in clinical trial settings?

Q3: Does CHMP agree with the proposed specific risk profiles for kidney transplantation which determine background risk of rejection associated with immunosuppressive therapy?

Q4: Does CHMP agree that long-term outcome after kidney transplantation is an area of unmet medical need, for which conditional marketing authorization procedures should be considered, to facilitate timely access to new therapies? If so, does CHMP agree with the proposed surrogate endpoints for clinical trials for therapies requiring conditional marketing authorization?

Q5: Does CHMP agree with the proposed patient reported outcomes as (primary/secondary) endpoints for use in clinical trials of kidney transplantation interventions?

transplantation (not specific for kidney transplantation) as a composite of four outcomes:

- Patient death
- Graft failure—defined by discrete criteria (e.g., permanent return to pre-transplantation treatment modality for a specific period)
- Biopsy-proven acute rejection—including pathological grading for the transplant, outcome, treatment, and response
- Graft (dys)function—defined by best available clear-cut and discrete criteria for kidney, lung, and heart transplantations (e.g., measurement of creatinine/inulin clearance for kidney dysfunction).

Components of the composite endpoint could be omitted for several reasons, such as if there is limited sensitivity/specificity for available biomarkers, or limited consensus about the importance of individual risk factors or cut-off values. Factors such as previous early graft loss (because of immunological factors), re-transplantation, human leukocyte antigen (HLA) mismatch, and presence of HLA antibodies are often taken into consideration, depending on the type of transplantation. The guideline also notes that best attempts should be undertaken to define the recipient's immunological risk at baseline, using categories such as "low/medium/high" or "elevated/non-elevated." Finally, CHMP/EWP/263148/06 notes that transplantation outcome is also influenced by surgery and comorbidity [10]. Therefore, reasonably validated scales for assessment of global transplantation risk are important and should be reflected in the target population of clinical studies.

## RATIONALE FOR UPDATING THE GUIDELINES

Not only has clinical organ transplantation changed markedly since the publication of CHMP/EWP/263148/06 [10]; the transplant community now observes signs of substantial decline in the rate of clinical innovation in this field. Consequently, ESOT—the umbrella organization under which all European transplant activities are organized—sought to understand whether updating guidelines on clinical trial endpoints might help to encourage kidney transplantation research, since potentially outdated or unclear definitions of risk groups and markers of success or failure might limit investment or innovation in transplantation medicine.

## METHODS

The project that ultimately resulted in the Broad Scientific Advice request and the present special issue began in May 2016, when the European Medicines Agency responded positively to ESOT's request to begin interactions relating to the overall topic. Within ESOT, the project was coordinated from its inception by Professor Maarten Naesens of KU Leuven, Belgium, with contributions by an international group of experts. This panel of volunteers included nephrologists, surgeons, transplant pathologists, epidemiologists, immunologists, and researchers (**Box 1**). The key events in the development process were:

- September 26, 2017: First workshop meeting at ESOT Barcelona to outline the project, define the core questions, and appoint working group leaders
- 2017–2018: Briefing package for the EMA was prepared by attendees of the workshop in Barcelona, who collaborated *via* telephone/e-mail and personal interaction
- June 11, 2018: Submission of a briefing package to EMA, outlining the scope and core questions put forward by ESOT
- June 20, 2018: Response from EMA and invitation to submit a formal request for Broad Scientific Advice
- 2018–2019: Establishment of five working groups related to the core questions agreed with EMA (**Box 2**), on which ESOT sought advice
- 2018–2019: Writing of the first drafts of the answers to the questions and searching consensus within the working groups
  - Each working group approached the Centre for Evidence in Transplantation (CET) with specific data extraction requests
  - Information provided by CET was used by each working group at their own discretion to produce draft documents
  - The documents were drafted by the experts within each working group
- September 17, 2019: Workshop at the ESOT congress in Copenhagen: consensus meeting to discuss the conclusions reached by the working groups
- 2019–2020: Finalization of the consensus document and preparation of the official request for Broad Scientific Advice

- June 5, 2020: Submission of the package to request Broad Scientific Advice from CHMP
- June 6–11, 2020: Start of the Scientific Advice Working Party (SAWP) procedure
- July 6–9, 2020: SAWP discussion meeting, at which a list of issues to be addressed by ESOT was adopted
- September 24, 2020: ESOT submitted additional documentation to the SAWP
- September 30, 2020: Virtual discussion meeting between ESOT and the SAWP, addressing the list of issues
- September 28 to October 1, 2020: SAWP agreed on the advice to be given to ESOT
- December 7–10, 2020: adoption of the advice by CHMP and response received by EMA
- 2020–2021: Reformatting of the package submitted to CHMP to fit publication style (in article form and as a positioning paper), to allow widespread dissemination of ESOT's answers to the questions and the CHMP advice
- 2021: The ESOT position on each question, including any comments from EMA, is summarized in evidence-based articles within this special issue.

## CONCLUSION

In June 2020, ESOT presented an extended form of the articles and supporting materials in this special issue to the EMA as one document, as part of the package to request Broad Scientific Advice from CHMP. Following constructive responses from the EMA, this special issue is published to communicate outcomes and extend discussions among the wider kidney transplant community. Ultimately, it is hoped that endpoints suitable for future clinical trials and better consensus on risk stratification are developed and agreed globally, thereby increasing the quality of future research and evidence, and advancing the practical management of kidney transplantation, particularly for long-term outcomes.

## AUTHOR CONTRIBUTIONS

This article is one of a series of papers developed from content relating to the Broad Scientific Advice request, submitted to the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) by the European Society for Organ Transplantation (ESOT) in 2020: interactions between the EMA and ESOT regarding this request began in 2016. The present article was adapted by MN from the final Broad Scientific Advice request submission (June 2020), presentation documents and minutes of the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020), to form an introduction to the special issue and present the rationale and methodology used throughout the process; the article was reviewed by SS and finalized by both co-authors.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

The authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. Coemans M, Süsal C, Döhler B, Anglicheau D, Giral M, Bestard O, et al. Analyses of the Short- and Long-Term Graft Survival after Kidney Transplantation in Europe between 1986 and 2015. *Kidney Int* (2018) 94: 964–73. doi:10.1016/j.kint.2018.05.018
2. O'Connell PJ, Kuypers DR, Mannon RB, Abecassis M, Chadban SJ, Gill JS, et al. Clinical Trials for Immunosuppression in Transplantation: the Case for Reform and Change in Direction. *Transplantation* (2017) 101:1527–34. doi:10.1097/TP.0000000000001648
3. Stegall MD, Morris RE, Alloway RR, Mannon RB. Developing New Immunosuppression for the Next Generation of Transplant Recipients: the Path Forward. *Am J Transplant* (2016) 16:1094–101. doi:10.1111/ajt.13582
4. Mayrdorfer M, Liefeldt L, Wu K, Rudolph B, Zhang Q, Friedersdorff F, et al. Exploring the Complexity of Death-Censored Kidney Allograft Failure. *J Am Soc Nephrol* (2021) 32:1513–26. doi:10.1681/asn.2020081215
5. Ying T, Shi B, Kelly PJ, Pilmore H, Clayton PA, Chadban SJ, et al. Death after Kidney Transplantation: an Analysis by Era and Time post-transplant. *J Am Soc Nephrol* (2020) 31:2887–99. doi:10.1681/asn.2020050566
6. Vanholder R, Domínguez-Gil B, Busic M, Cortez-Pinto H, Craig JC, Jager KJ, et al. Organ Donation and Transplantation: a Multi-Stakeholder Call to Action. *Nat Rev Nephrol* (2021) 17:554–68. doi:10.1038/s41581-021-00425-3
7. Abramowicz D, Oberbauer R, Heemann U, Viklicky O, Peruzzi L, Mariat C, et al. Recent Advances in Kidney Transplantation: a Viewpoint from the Descartes Advisory Board*. *Nephrol Dial Transplant* (2018) 33:1699–707. doi:10.1093/ndt/gfx365

8. Naesens M, Thaunat O. BENEFIT of Belatacept: Kidney Transplantation Moves Forward. *Nat Rev Nephrol* (2016) 12:261–2. doi:10.1038/nrneph.2016.34
9. Bestard O, Couzi L, Crespo M, Kessaris N, Thaunat O. Stratifying the Humoral Risk of Candidates to a Solid Organ Transplantation: a Proposal of the ENGAGE Working Group. *Transpl Int* (2021) 34:1005–18. doi:10.1111/tri.13874
10. European Medicines Agency. Clinical Investigation of Immunosuppressants for Solid Organ Transplantation. CHMP/EWP/263148/06 (2008). Available at: https://www.ema.europa.eu/en/clinical-investigation-immunosuppressants-solid-organ-transplantation (Accessed October 18, 2021).

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10141

# Evolution of the Definition of Rejection in Kidney Transplantation and Its Use as an Endpoint in Clinical Trials

*Jan Ulrich Becker*[1†], *Daniel Seron*[2†], *Marion Rabant*[3], *Candice Roufosse*[4] and *Maarten Naesens*[5*]

[1]Institute of Pathology, University Hospital Cologne, Cologne, Germany, [2]Department of Nephrology and Kidney Transplantation, Vall d'Hebrón University Hospital, Barcelona, Spain, [3]Department of Pathology, Hôpital Necker–Enfants Malades, Paris, France, [4]Centre for Inflammatory Disease, Department of Immunology and Inflammation, Imperial College London, London, United Kingdom, [5]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

This article outlines the evolving definition of rejection following kidney transplantation. The viewpoints and evidence presented were included in documentation prepared for a Broad Scientific Advice request to the European Medicines Agency (EMA), relating to clinical trial endpoints in kidney transplantation. This request was initiated by the European Society for Organ Transplantation (ESOT) in 2016 and finalized following discussions between the EMA and ESOT in 2020. In ESOT's opinion, the use of "biopsy-proven acute rejection" as an endpoint for clinical trials in kidney transplantation is no longer accurate, although it is still the approved histopathological endpoint. The spectrum of rejection is now divided into the phenotypes of borderline changes, T cell-mediated rejection, and antibody-mediated rejection, with the latter two phenotypes having further subclassifications. Rejection is also described in relation to graft (dys)function, diagnosed because of protocol (surveillance) or indication (for-cause) biopsies. The ongoing use of outdated terminology has become a potential barrier to clinical research in kidney transplantation. This article presents these perspectives and issues, and provides a foundation on which subsequent articles within this Special Issue of Transplant International build.

**Keywords: biopsy, subclinical rejection, antibody-mediated rejection, T cell-mediated rejection, borderline changes, kidney transplantation outcome**



***Correspondence:**
Maarten Naesens
maarten.naesens@kuleuven.be

[†]These authors have contributed equally to this work

***Received:*** *21 October 2021*
***Accepted:*** *11 January 2022*
***Published:*** *20 May 2022*

***Citation:***
*Becker JU, Seron D, Rabant M, Roufosse C and Naesens M (2022) Evolution of the Definition of Rejection in Kidney Transplantation and Its Use as an Endpoint in Clinical Trials. Transpl Int 35:10141. doi: 10.3389/ti.2022.10141*

## INTRODUCTION

The approved histopathological endpoint for clinical trials of kidney transplantation is the presence or absence of biopsy-proven acute rejection (BPAR) (1). This endpoint has not changed for decades, despite many improvements in diagnostic assessment, immunosuppression, and monitoring protocols for kidney transplant recipients, as well as developments in our understanding of the epidemiology and pathophysiology of rejection (2).

Over time, the spectrum of rejection has broadened, with distinctions made between two major subtypes: T cell-mediated rejection (TCMR) and antibody-mediated rejection (AMR) (3). Deeper distinctions have also been made between acute (or active) and chronic phenotypes of TCMR and AMR, as defined in the Banff Classification (2), and subtypes within these phenotypes. In addition, evidence has emerged to indicate that non-specific acute rejection, or early TCMR, is becoming less relevant as the primary endpoint in kidney transplantation (4) because it is no longer considered a strong predictor of graft loss. Ongoing use of outdated terminology and definitions of

histopathological endpoints such as BPAR in clinical trials has therefore become a potential barrier to research, particularly for drug development programs that aim specifically at treating only one main rejection subtype.

Furthermore, the strategy of performing protocol biopsies in the early years following transplantation has been adopted by several European centers, to detect subclinical rejection and guide ongoing patient management (5). It has become important, therefore, to consider whether endpoints defined for indication biopsies are also valid for protocol biopsies.

## REJECTION PHENOTYPES

The classification of allograft rejection has often been modified over the years, such that six histological rejection phenotypes are widely described (2, 6):

- Suspicious (borderline) for acute TCMR (henceforth simplified to "borderline changes")
- Acute TCMR (aTCMR; classified as IA, IB, IIA, IIB, III)
- Chronic active TCMR (caTCMR)
- Acute/active antibody-mediated rejection (aAMR)
- Chronic antibody-mediated rejection (cAMR)
- Chronic active antibody-mediated rejection (caAMR).

Borderline changes represent less severe inflammation scores than aTCMR. The threshold of inflammation used for diagnosis of borderline changes (interstitial inflammation [i]0, <10% of the non-fibrotic cortex; or i1, 10%–25% of the non-fibrotic cortex) varies among centers, because between 2005 and 2017 the Banff Classification stated that retaining the i1 threshold for borderline changes with tubulitis (t) > 0 was permitted (7). However, in 2019 the minimal threshold changed to i1t1, given that several studies indicated that isolated tubulitis in the absence of interstitial inflammation (i0) did not associate with impaired graft outcome—a finding supported by most of those involved in ratifying the Banff 2019 update (7-11). In addition, decreased heterogeneity in center practice is anticipated (11). Banff 2019 also emphasized that "borderline changes" should be known as "borderline (suspicious) for acute TCMR," to make a clear reference to rejection and treatment (11).

In the 1990s, a diagnosis of aTCMR was based on a clinical definition (i.e., an acute rise of serum creatinine that responded to antirejection therapy) and/or a clinicopathological definition (i.e., acute rejection, being aTCMR or borderline changes in an indication biopsy) (12, 13). The criteria for aTCMR have not changed since the original 1997 Banff Classification and the scores remain based on the presence of interstitial inflammation (i), tubulitis (t), and arteritis (v). However, tubulitis is now considered in all but severely atrophic cortical tubules as either Banff lesion score t or t-IFTA (defined below), whereas previously it was only considered in mildly atrophic or non-atrophic tubules (11).

The impact of inflammation in atrophic areas (i-IFTA) on graft outcomes has been widely demonstrated (8, 14–16), and the effect of i-IFTA on graft survival was not significantly affected by treatment for concomitant aTCMR (15); i-IFTA has also been shown to be related to under-immunosuppression and is more commonly preceded by aTCMR than biopsies without i-IFTA (16, 17), although in some reports the majority of cases with i-IFTA did not have a previous biopsy with rejection (18). These findings suggest that i-IFTA could partly reflect alloimmunity, although further research is warranted. The same applies for tubulitis in moderately atrophic tubules captured as Banff lesion Score t-IFTA (16).

The Banff 2015 meeting noted for the first time that caTCMR could manifest in tubulointerstitial and vascular compartments, and at the 2017 meeting the proposal to include inflammation in areas of fibrosis was incorporated into the consensus classification as caTCMR (2). This classification requires interstitial inflammation involving >25% of the total cortex (ti score 2 or 3) and >25% of the sclerotic cortical parenchyma (i-IFTA score 2 or 3) with moderate tubulitis (t2) involving one or more tubules, not including severely atrophic tubules, while other known causes of i-IFTA are ruled out. Excluding other causes of inflammation in fibrosed areas is important, as i-IFTA is not a specific lesion and can be seen in cases of polyomavirus infection, pyelonephritis, AMR, recurrent glomerulonephritis, and obstruction. Inflammation might instead be an indication of very recent nephron loss as consequence (rather than the cause) of the injury *per se*. The response of caTCMR to increased doses of immunosuppressive therapy has not been studied (2).

In 2001, specific criteria for AMR were introduced (3), linking histopathological changes, presence of C4d, and presence of donor-specific antibodies (DSA). These were revised in 2007 (19) with the introduction of peritubular capillary (PTC) and C4d scores, and cAMR. In 2013, C4d-negative AMR was recognized, and C4d was replaced by a sign of interaction between the DSA and the endothelium (positive C4d or microcirculation inflammation, glomerulitis and peritubular capillaritis [g + PTC] ≥2, or molecular markers) (20). Finally, and importantly, in 2017 the classification for AMR was revised a second time, with acceptance of positive C4d staining as substitute for DSA in the serological criterion for DSA-negative cases and elimination of the suspicious for AMR category (not fulfilling all three criteria). Criteria for AMR were unchanged in 2019.

In addition, rejection phenotypes of kidney transplants are distinguished according to their association with graft (dys)function. Protocol (surveillance) biopsies are performed, per definition, at the time of stable graft function to detect subclinical inflammation (subclinical aTCMR and AMR) (5). Indication (for-cause) biopsies are performed at the time of graft dysfunction.

Finally, although molecular diagnostics of kidney transplant rejection has been validated prospectively in a multicentric fashion (21) and is currently applied for secondary endpoints in clinical trials, we do not consider mRNA expression patterns a valid primary endpoint at this time. Banff has not formally recognized this particular assay and is moving towards an entirely different technological platform (22) which will also

need rigorous validation for diagnostic or theranostics use, before being proposed as primary endpoint for clinical trials.

## CONCLUSIONS

ESOT has come to the following conclusions:

- The use of BPAR as an endpoint for clinical trials in kidney transplantation is no longer accurate.
  - Using outdated and/or non-specific definitions, such as BPAR, compromises the future of high-quality clinical research, especially for interventions that are targeted at one rejection subtype.
- Kidney transplant rejection should be classified by its phenotypes—borderline changes, TCMR, and AMR (the two latter having subtypes), and in relation to the nature of graft (dys)function (i.e., indication [for-cause] vs. protocol [surveillance] biopsies).

### Scientific Advice From the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA) About These Conclusions

- The CHMP acknowledged that histological subclassifications of rejection have evolved during the last decade.
- The CHMP agreed that the histological subtype of rejection is a useful specification and noted that this detailing might be very informative in profiling efficacy of immunosuppression.
- The CHMP commented that the reason for undertaking a protocol or indication biopsy should be taken into consideration when defining endpoints for clinical trials.

## AUTHOR CONTRIBUTIONS

This article is one of several papers developed from the Broad Scientific Advice request, submitted to the EMA/CHMP by ESOT in 2020: interactions between the EMA and ESOT regarding this request began in 2016. For the present article, working groups on histological and functional endpoints in kidney transplantation developed the ESOT position on the core question "Does CHMP agree with the updated definitions of rejection and their potential use as primary endpoints in studies of kidney transplantation?". The Centre for Evidence in Transplantation provided support with specific data extractions: these literature searches formed the basis of evidence used in the advice request and present article. Input into the working groups was provided from all ESOT members involved in the advice request process. The present article was adapted by MN from the Broad Scientific Advice request submission documents and minutes of the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP), and the final response from the SAWP (December 2020). The article was revised by JUB and DS, and circulated to MR, CR, and MN for e-mail review. The article was finalized and approved by all co-authors before submission.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

JUB consults for Sanofi. MR has received lecture fees from Astellas and Chiesi; and research grant support (paid to institution) from Astellas and Chiesi for investigator-initiated studies.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Literature searches were undertaken by Liset Pengel, Centre for Evidence in Transplantation, Oxford, United Kingdom. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. European Medicines Agency. *Clinical Investigation of Immunosuppressants for Solid Organ Transplantation* (2008). CHMP/EWP/263148/06.
2. Haas M, Loupy A, Lefaucheur C, Roufosse C, Glotz D, Seron D, et al. The Banff 2017 Kidney Meeting Report: Revised Diagnostic Criteria for Chronic Active T Cell-Mediated Rejection, Antibody-Mediated Rejection, and Prospects for Integrative Endpoints for Next-Generation Clinical Trials. *Am J Transplant* (2018) 18(2):293–307. doi:10.1111/ajt.14625
3. Racusen LC, Colvin RB, Solez K, Mihatsch MJ, Halloran PF, Campbell PM, et al. Antibody-Mediated Rejection Criteria - an Addition to the Banff '97 Classification of Renal Allograft Rejection. *Am J Transplant* (2003) 3:708–14. doi:10.1034/j.1600-6143.2003.00072.x
4. Klintmalm GB, Vincenti F, Kirk A Steroid-Responsive Acute Rejection Should Not Be the End Point for Immunosuppressive Trials. *Am J Transplant* (2016) 16:3077–8. doi:10.1111/ajt.13889
5. Chapman JR Do protocol Transplant Biopsies Improve Kidney Transplant Outcomes? *Curr Opin Nephrol Hypertens* (2012) 21:580–6. doi:10.1097/mnh.0b013e32835903f4

6. Roufosse C, Simmonds N, Groningen MC, Haas M, Henriksen KJ, Horsfield C, et al. Reference Guide to the Banff Classification of Renal Allograft Pathology. *Transplantation* (2018) 102:1795–814. doi:10.1097/TP.0000000000002366
7. Becker JU, Chang A, Nickeleit V, Randhawa P, Roufosse C. Banff Borderline Changes Suspicious for Acute T-Cell Mediated Rejection: where Do We Stand? *Am J Transplant* (2016) 16(9):2654–60. doi:10.1111/ajt.13784
8. Nankivell BJ, P'Ng CH, Chapman JR. Does Tubulitis without Interstitial Inflammation Represent Borderline Acute T Cell Mediated Rejection? *Am J Transplant* (2019) 19(1):132–44. doi:10.1111/ajt.14888
9. Wiebe C, Rush DN, Gibson IW, Pochinco D, Birk PE, Goldberg A, et al. Evidence for the Alloimmune Basis and Prognostic Significance of Borderline T Cell-Mediated Rejection. *Am J Transplant* (2020) 20:2499–508. doi:10.1111/ajt.15860
10. McRae M, Bouchard-Boivin F, Béland S, Noël R, Côté I, Lapointe I, et al. Impact of the Current versus the Previous Diagnostic Threshold on the Outcome of Patients with Borderline Changes Suspicious for T Cell-Mediated Rejection Diagnosed on Indication Biopsies. *Transplantation* (2018) 102:2120–5. doi:10.1097/tp.0000000000002327
11. Loupy A, Haas M, Roufosse C, Maarten N, Benjamin A, Marjan A, et al. The Banff 2019 Kidney Meeting Report (I): Updates on and Clarification of Criteria for T Cell- and Antibody-Mediated Rejection. *Am J Transplant* (2020) 20:2318–31. doi:10.1111/ajt.15898
12. Meier-Kriesche H-U, Ojo AO, Hanson JA, Cibrik DM, Punch JD, Leichtman AB, et al. Increased Impact of Acute Rejection on Chronic Allograft Failure in Recent Era. *Transplantation* (2000) 70:1098–100. doi:10.1097/00007890-200010150-00018
13. Serón D, Arias M, Campistol JM, Morales JM. Late Renal Allograft Failure between 1990 and 1998 in Spain: a Changing Scenario. *Transplantation* (2003) 76:1588–94. doi:10.1097/01.TP.0000092495.07385.3C
14. Mengel M, Reeve J, Bunnag S, Einecke G, Jhangri GS, Sis B, et al. Scoring Total Inflammation Is superior to the Current Banff Inflammation Score in Predicting Outcome and the Degree of Molecular Disturbance in Renal Allografts. *Am J Transplant* (2009) 9:1859–67. doi:10.1111/j.1600-6143.2009.02727.x
15. Mannon RB, Matas AJ, Grande J, Leduc R, Connett J, Kasiske B, et al. Inflammation in Areas of Tubular Atrophy in Kidney Allograft Biopsies: a Potent Predictor of Allograft Failure. *Am J Transplant* (2010) 10:2066–73. doi:10.1111/j.1600-6143.2010.03240.x
16. Lefaucheur C, Gosset C, Rabant M, Viglietti D, Verine J, Aubert O, et al. T Cell-Mediated Rejection Is a Major Determinant of Inflammation in Scarred Areas in Kidney Allografts. *Am J Transplant* (2018) 18:377–90. doi:10.1111/ajt.14565
17. Nankivell BJ, Shingde M, Keung KL, Fung CLS, Borrows RJ, O'Connell PJ, et al. The Causes, Significance and Consequences of Inflammatory Fibrosis in Kidney Transplantation: The Banff i-IFTA Lesion. *Am J Transplant* (2018) 18:364–76. doi:10.1111/ajt.14609
18. Helgeson ES, Mannon R, Grande J, Gaston RS, Cecka MJ, Kasiske BL, et al. I-IFTA and Chronic Active T Cell-Mediated Rejection: A Tale of 2 (DeKAF) Cohorts. *Am J Transplant* (2021) 21:1866–77. doi:10.1111/ajt.16352
19. Solez K, Colvin RB, Racusen LC, Haas M, Sis B, Mengel M, et al. Banff 07 Classification of Renal Allograft Pathology: Updates and Future Directions. *Am J Transplant* (2008) 8:753–60. doi:10.1111/j.1600-6143.2008.02159.x
20. Haas M, Sis B, Racusen LC, Solez K, Glotz D, Colvin RB, et al. Banff 2013 Meeting Report: Inclusion of C4d-Negative Antibody-Mediated Rejection and Antibody-Associated Arterial Lesions. *Am J Transplant* (2014) 14:272–83. doi:10.1111/ajt.12590
21. Halloran PF, Reeve J, Akalin E, Aubert O, Bohmig GA, Brennan D, et al. Real Time Central Assessment of Kidney Transplant Indication Biopsies by Microarrays: The INTERCOMEX Study. *Am J Transplant* (2017) 17(11):2851–62. doi:10.1111/ajt.14329
22. Mengel M, Loupy A, Haas M, Candice R, Maarten N, Enver A, et al. Banff 2019 Meeting Report: Molecular Diagnostics in Solid Organ Transplantation - Consensus for the Banff Human Organ Transplant (B-HOT) Gene Panel and Open Source Multicenter Validation. *Am J Transplant* (2020) 20(9):2305–17. doi:10.1111/ajt.16059

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10135

# Proposed Definitions of T Cell-Mediated Rejection and Tubulointerstitial Inflammation as Clinical Trial Endpoints in Kidney Transplantation

*Daniel Seron[1†], Marion Rabant[2†], Jan Ulrich Becker[3], Candice Roufosse[4], Maria Irene Bellini[5], Georg A. Böhmig[6], Klemens Budde[7], Fritz Diekmann[8], Denis Glotz[9], Luuk Hilbrands[10], Alexandre Loupy[11], Rainer Oberbauer[12], Liset Pengel[13], Stefan Schneeberger[14] and Maarten Naesens[15]**

[1]Department of Nephrology and Kidney Transplantation, Vall d'Hebrón University Hospital, Barcelona, Spain, [2]Department of Pathology, Hôpital Necker–Enfants Malades, Paris, France, [3]Institute of Pathology, University Hospital Cologne, Cologne, Germany, [4]Centre for Inflammatory Disease, Department of Immunology and Inflammation, Imperial College London, London, United Kingdom, [5]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy, [6]Division of Nephrology and Dialysis, Department of Internal Medicine, Medical University of Vienna, Vienna, Austria, [7]Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany, [8]Department of Nephrology and Kidney Transplantation, Hospital Clinic Barcelona, Barcelona, Spain, [9]Paris Translational Research Center for Organ Transplantation, Hôpital Saint Louis, Paris, France, [10]Department of Nephrology, Radboud University Medical Center, Nijmegen, Netherlands, [11]Paris Translational Research Center for Organ Transplantation, Hôpital Necker, Paris, France, [12]Department of Nephrology and Dialysis, Medical University of Vienna, Vienna, Austria, [13]Centre for Evidence in Transplantation, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom, [14]Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria, [15]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

The diagnosis of acute T cell-mediated rejection (aTCMR) after kidney transplantation has considerable relevance for research purposes. Its definition is primarily based on tubulointerstitial inflammation and has changed little over time; aTCMR is therefore a suitable parameter for longitudinal data comparisons. In addition, because aTCMR is managed with antirejection therapies that carry additional risks, anxieties, and costs, it is a clinically meaningful endpoint for studies. This paper reviews the history and classifications of TCMR and characterizes its potential role in clinical trials: a role that largely depends on the nature of the biopsy taken (indication vs protocol), the level of inflammation observed (e.g., borderline changes vs full TCMR), concomitant chronic lesions (chronic active TCMR), and the therapeutic intervention planned. There is ongoing variability—and ambiguity—in clinical monitoring and management of TCMR. More research, to investigate the clinical relevance of borderline changes (especially in protocol biopsies) and effective therapeutic strategies that improve graft survival rates with minimal patient morbidity, is urgently required. The present paper was developed from documentation produced by the European Society for Organ Transplantation (ESOT) as part of a Broad Scientific Advice request that ESOT submitted to the European Medicines Agency for discussion in 2020. This paper proposes to move toward refined definitions of aTCMR and borderline changes to be included as primary endpoints in clinical trials of kidney transplantation.

**Keywords: kidney transplantation, outcomes, EMA guideline, T cell-mediated rejection, borderline changes**

OPEN ACCESS

***Correspondence:**
Maarten Naesens
maarten.naesens@kuleuven.be

†These authors have contributed equally to this work

**Received:** 21 October 2021
**Accepted:** 11 January 2022
**Published:** 20 May 2022

**Citation:**
Seron D, Rabant M, Becker JU, Roufosse C, Bellini MI, Böhmig GA, Budde K, Diekmann F, Glotz D, Hilbrands L, Loupy A, Oberbauer R, Pengel L, Schneeberger S and Naesens M (2022) Proposed Definitions of T Cell-Mediated Rejection and Tubulointerstitial Inflammation as Clinical Trial Endpoints in Kidney Transplantation. Transpl Int 35:10135. doi: 10.3389/ti.2022.10135

## ACUTE T CELL-MEDIATED REJECTION ENDPOINTS: THE HISTORY

Health authorities have long accepted biopsy-proven acute rejection (BPAR) as a primary efficacy variable in clinical trials for the prevention and treatment of transplant rejection (1): in epidemiological studies performed during the 1990s (2, 3), BPAR was associated with poor long-term outcomes. There is a general belief that BPAR is considered to reflect acute T cell-mediated rejection (aTCMR), likely in part related to the fact that between 1991 and 2001 the recognition of antibody-mediated rejection (AMR) was limited to its hyperacute/accelerated forms and overshadowed by grading of aTCMR in the Banff Classification for Allograft Pathology (4). This belief extends to the fact that many pivotal studies of immunosuppressant therapy have utilized the term "BPAR" to describe what was often more specifically aTCMR, identified on indication biopsies (discussed below). Indeed, the opinion that BPAR and aTCMR are interchangeable terms remains largely speculative; evidence indicates that they are not equal, if only because the definition of BPAR does not discriminate between the different subtypes of rejection that have been identified (see Becker et al. (5), this special issue).

However, as AMR was only introduced into the Banff Classification later (4), and as the specific definition of aTCMR has not changed markedly since 1997, with some caveats the aTCMR diagnosis can likely be used for between-study comparisons over time. The fact that aTCMR is managed with antirejection therapies that cause risk, anxiety, and cost continues to make aTCMR a clinically relevant endpoint for research purposes. TCMR was found to be an important cause for graft failure in a recent retrospective study (6). Of note, interobserver variability in the diagnosis of aTCMR is high (7), which warrants caution in the interpretation of single-center data without central pathological review.

In terms of drug development studies in kidney transplantation, European Medicines Agency and US Food & Drug Administration approvals of mycophenolate mofetil, daclizumab, tacrolimus, basiliximab, and sirolimus were based primarily on superiority findings, with BPAR (more specifically, rates of aTCMR in indication biopsies) included as the primary efficacy variable or part of a composite measure, with graft failure and patient death (8–11). In the Symphony study (12), 1,645 kidney transplant recipients were randomized to combination therapy involving mycophenolate mofetil and corticosteroids, with or without cyclosporine, daclizumab induction, tacrolimus, or sirolimus; kidney function (evaluated by estimated [e] glomerular filtration rate [GFR] at 1 year post transplantation) was the primary efficacy variable and BPAR was the secondary efficacy variable. Kidney function and graft survival rates were better in tacrolimus-treated patients compared with others: the BPAR rate (of unspecified subtype; presumably mostly aTCMR) was lowest in those receiving low-dose tacrolimus (12%) compared with standard-dose cyclosporine (26%), low-dose cyclosporine (24%), or low-dose sirolimus (37%) (12).

The Symphony trial therefore defined a new standard of care in kidney transplantation that was widely employed thereafter because of its efficacy in preserving function and preventing rejection. Subsequent studies also reported similarly low rates of BPAR for innovative combination regimens (13–17). Collectively, this research showed that the incidence of BPAR in indication biopsies could be modulated by immunosuppressive therapy and has decreased considerably over time, leading to improvements in post-transplantation treatment and understanding of acute rejection (18).

## ACUTE TCMR IN INDICATION BIOPSIES IN SUPERIORITY OR NON-INFERIORITY STUDIES

Now that the incidence of aTCMR is consistently reported at ~10% during the first year following kidney transplantation (19–21), it is important to reconsider its utility as a primary efficacy variable. The low prevalence of aTCMR with current immunosuppressive regimens, and the less consistent association of aTCMR with outcome (22–26), indicate a limited need for superiority trials that aim to further reduce rates of aTCMR. Any benefits gained from such trials would be outweighed by the considerable drawbacks associated with powerful regimens that risk over-immunosuppression and create safety or tolerability issues for many patients. Nevertheless, including biopsy-proven aTCMR as a primary efficacy variable in non-inferiority trials remains highly relevant, since aTCMR in an indication biopsy leads to heightened therapeutic interventions, treatment burden, morbidity, and cost. Treatment-resistant TCMR may also lead to graft loss, or in less severe cases to nephron loss, with detrimental long-term consequences for graft function.

Definition and presentation of the rejection subtype, and its association with outcome, are important considerations for discussions exploring the value of aTCMR as a primary efficacy variable in clinical trials. For example, AMR was not clearly defined until 2001 (4): it is likely that some patients considered to have BPAR in the 1990s might have experienced an unrecognized episode of AMR or mixed AMR–aTCMR. Consequently, the relationship between BPAR (i.e., aTCMR) and outcome may have been overemphasized in the past.

In addition to the rejection subtype, one can also reflect on how characteristics of donors and recipients, and the incidence of rejection, have changed in a time-dependent manner with emerging evidence and improvements in practice. Studies evaluating the relationship between aTCMR and graft survival have therefore yielded seemingly contradictory results (22–25): they indicate that further reductions of the incidence of aTCMR will not directly translate into better rates of long-term graft survival, and also suggest that higher incidences of aTCMR do not correlate with incidences of graft failure.

For example, in an epidemiological study that distinguished between aTCMR and AMR, aTCMR diagnosed by indication biopsies was not associated with decreased graft survival rates (22). In the Tricontinental Mycophenolate Mofetil Renal Transplantation Study, outcome evaluation at 3 years (i.e. 3-year graft survival rate) did not show any benefit for cyclosporine plus mycophenolate mofetil over cyclosporine

plus azathioprine, despite a significant reduction in the incidence of rejection during the first year. However, this study was not adequately powered to detect a difference in 3-year graft survival rates (23). Of note, patients included in the Tricontinental study in Australia were followed for 15 years; again, no long-term benefit of mycophenolate mofetil was observed (24).

Conversely, in a 5-year follow-up of a study comparing steroid continuation or withdrawal in a tacrolimus plus mycophenolate mofetil-based regimen, the acute rejection rate increased after steroid withdrawal and was associated with decreased survival (25). Analysis of the Australian and New Zealand Dialysis and Transplant Registry (13,614 recipients) showed that aTCMR was associated with allograft survival and death with a functioning graft (specifically, death due to cardiovascular disease or cancer) (26).

Similarly, in the belatacept trial (27), more-intensive and less-intensive belatacept regimens were compared with a cyclosporine-based regimen, with BPAR, graft loss, and recipient death as the composite primary endpoint that was used to demonstrate non-inferiority, and GFR as the endpoint to show superiority. Despite higher incidence of BPAR during the first year (22% in the more-intensive belatacept group, 17% in the less-intensive belatacept group, 7% in the cyclosporine group), kidney function and long-term allograft survival were superior in patients receiving belatacept. However, when belatacept- or cyclosporine-treated recipients with or without rejection were compared, GFR was significantly lower in those who experienced an episode of acute rejection, suggesting nephron loss in patients experiencing BPAR (or aTCMR). This supports the prognostic meaning of rejection, even in patients on belatacept. The proportion of patients who developed *de novo* (*dn*) donor-specific antibodies (DSA) at 7 years was decreased in belatacept-treated patients compared with cyclosporine-treated patients (28), illustrating that most BPAR cases were aTCMR, and that a higher rate of aTCMR did not translate into a worse outcome in this trial. The poorer kidney function in cyclosporine-compared with belatacept-treated patients could be explained by nephrotoxicity, not rejection, although rejection still affected graft function within the belatacept arm. It is unclear whether lower *dn*DSA, despite higher aTCMR, could be explained by a specific effect of belatacept, or better adherence to this treatment compared with cyclosporine.

Mixed results on the association between aTCMR and outcome were corroborated by indication-biopsy findings reported for 256 kidney transplant recipients with aTCMR, treated with steroids (29). Overall graft survival rates were 85% after 5 years and 69% after 10 years. Best predictors of allograft loss were GFR, presence of inflammation in areas of interstitial fibrosis and tubular atrophy (i-IFTA) on 3-month protocol biopsies, and presence of anti-HLA (human leukocyte antigen) DSAs at 3 months. This suggests that transition from aTCMR to chronic active (ca)TCMR or response to aTCMR treatment constitute hallmarks of poor long-term outcome; it also illustrates that not all aTCMR episodes are equal, at least in terms of their treatment response. For example, patients with aTCMR (on indication biopsy) and a GFR >44 ml/min, no or mild i-IFTA, and no anti-HLA-DSA had a 74% graft survival rate at 10 years, whereas those with aTCMR and i-IFTA grade 2 or 3 had a 55% graft survival rate at this time point (29). It is obvious that aTCMR may cause injury to the nephron, which might result in subsequent nephron loss, as evidenced by the fact that aTCMR contributed to graft loss in ~34% of failures (6). Older data also suggested that aTCMR grade II [with intimal arteritis) conferred less responsiveness to steroid therapy and a poorer prognosis for graft survival than grade 1 aTCMR (30, 31)]. Notwithstanding these data, more research is needed to better define the aTCMR phenotypes that confer increased risk of worse outcome, which is potentially of importance for the choice of the primary efficacy variable in future clinical trials.

## BORDERLINE CHANGES IN INDICATION BIOPSIES

As the presence of aTCMR in indication biopsies is perhaps less important than it was in the early years of kidney transplantation, the relevance of borderline changes in such biopsies might be even more trivial, given that these represent less severe inflammation scores than aTCMR. Nevertheless, sampling errors and low reproducibility of Banff Lesion Scores could lead to arbitrary classifications, and less strict distinctions between aTCMR and borderline changes that do not reach the aTCMR threshold. This is corroborated by molecular analysis of biopsies showing borderline changes at the time of graft dysfunction, which illustrates that such changes represent a molecularly heterogeneous group: some do not resemble rejection, whereas others resemble aTCMR (32).

Indication biopsies are undertaken when there are clinical signs of deteriorating kidney function. Although borderline changes detected on indication biopsies are less likely to be associated with graft failure than aTCMR, 50–80% of cases of borderline changes detected in indication biopsies receive antirejection treatment with high-dose corticosteroids (33–35). Consequently, borderline changes can be clinically relevant even in the absence of more severe lesions, because of the impact of any decision to initiate antirejection therapy (36).

Despite this clinical relevance, the association between borderline changes in indication biopsies and graft outcome has not been widely studied in the current context of transplantation medicine; the limited research findings are mixed. A retrospective analysis illustrated that graft survival rates were significantly better in patients with borderline changes than in those with aTCMR, but significantly worse than in the control group, despite antirejection therapy, similar clinical characteristics, and similar graft dysfunction at time of biopsy (37). More recently, comparison of patients with different lesion scores for borderline rejection showed that the occurrence of death-censored graft failure or doubling of serum creatinine concentration post biopsy at 5 years was 5% for those scoring t1i0 but reached 14% for those scoring ≥t1i1. These endpoints also occurred in 5% of recipients with no rejection and 21% of those with TCMR. Patients with biopsy lesion scores of t1i0 therefore had a prognosis similar to that of non-rejectors (adjusted hazard ratio [HR] 0.6; 95% confidence interval [CI]

0.1–2.2), and better than that of patients with lesions scoring ≥t1i1 (adjusted HR 3.8; 95% CI 1.3–11.5) (4).

In a study of 803 renal transplantations, Wiebe et al. found an independent correlation between HLA-DR/DQ molecular mismatch, presence of borderline changes (diagnosed according to the Banff 1997 definition), and severity and frequency of rejection episodes. These investigators suggested that borderline changes could be part of a spectrum of alloimmune-mediated inflammation, not simply a response to injury (38).

The place of borderline changes as an endpoint in clinical trials is therefore not entirely clear, but there is evidence of its clinical relevance as far as observed in indication biopsies. Many registration studies for immunosuppressant therapies in kidney transplantation that utilized BPAR as the endpoint did not specify either the grade of rejection or the inclusion of borderline changes (8, 39–45). In registration studies for basiliximab (46,47) and belatacept (48), the definition of BPAR excluded borderline changes; only grades I or II aTCMR were considered in the BPAR definition (49). Only the ZEUS trial included borderline changes in its BPAR definition (49).

In Wu et al.'s retrospective analysis, borderline changes were treated with antirejection drugs, leading to complete reversibility in 57%, partial reversibility in 39%, and no reversibility in only 4% of cases (vs. 15% and 21% no reversibility for TCMR grades I and II, respectively) (37). Similarly, an earlier and smaller retrospective study (25) reported a high likelihood of complete response with antirejection treatment for borderline changes. Finally, in another retrospective study, outcome after determination of borderline changes (by serum creatinine and/or subsequent histology) showed that untreated changes were non-progressive in 72% of cases (50), although some biopsies were performed per protocol and thus not considered indication biopsies. This study suggested that conservative management of borderline changes in indication biopsies, at least in the short term, might be more appropriate than routine treatment as indicated for acute rejection.

As these were retrospective studies with mixed results, no conclusions can be drawn on the necessity or timing of any treatment for borderline changes in indication biopsies. The decision depends on center practice and clinician's judgment. However, the participants at the Banff 2019 meeting agreed that any findings below the i1t1 threshold would not be considered borderline changes as they are not associated with impaired graft outcome. In addition, as antirejection therapy is associated with treatment burden, comorbidity, anxiety, and heightened cost, a diagnosis of borderline changes that leads to therapeutic intervention represents a clinically impactful event. In a clinical trial setting, this could be a relevant marker for evaluating non-inferiority, despite having a limited association with graft failure and higher likelihood of reversibility on treatment, compared with aTCMR.

## BORDERLINE CHANGES OR TCMR IN PROTOCOL BIOPSIES

The potential utility of protocol biopsy histology as a primary efficacy variable has been evaluated in clinical trials of interventions that aim to prevent subclinical inflammation during the first year following kidney transplantation. Here, we consider the association between subclinical inflammation and outcome in patients not treated or treated for this condition, and the effect of basal immunosuppression on the incidence of subclinical inflammation. Since there is no international consensus on the definition of protocol biopsies, and the definitions used are not always explicitly stated in papers, there is some heterogeneity in the literature. Some centers define protocol biopsies according to the prescheduled nature of the biopsy; others take graft functional characteristics into account. For future clinical trials, the term "protocol biopsy" should be defined precisely, in terms of allowed change in serum creatinine or proteinuria, to improve interpretability of the results. International standardization of the definition of protocol biopsy would be highly welcomed.

Subclinical rejection (and/or borderline changes), identified in protocol biopsies during the first year post transplantation, are associated with progression of IFTA and increased serum creatinine levels (51, 52), impaired glomerular adaptation (53, 54), *dn*DSA appearance (52–56), and decreased graft survival (57).

Presence of borderline changes alone is associated with persistent inflammation in serial protocol biopsies, IFTA progression (58), *dn*DSA appearance, and decreased graft survival 51]. Nankivell et al. compared 146 patients with borderline changes (92 subclinical and 54 clinical episodes) versus 826 normal controls and 55 aTCMR patients. Subclinical borderline changes improved on subsequent protocol biopsies in 72% of cases but persisted in 19% and worsened in 9%. Untreated subclinical borderline changes resolved in 62% of cases, persisted in 27%, and worsened in 12%. Overall, presence of borderline changes remained an independent predictor of graft failure when adjusted for multiple immunological risk factors, time since transplant, and biopsy indication (51). However, the retrospective and associative nature of these data is a clear limitation of these studies, and bias introduced by attending physicians in treatment decision-making means the findings should be interpreted cautiously.

In a large study of 1-year protocol biopsies conducted in patients transplanted between 2000 and 2010, 73% of patients did not show rejection (with borderline changes counted as no rejection), 13% showed aTCMR ($i \geq 2$ and $t \geq 2$), and 14% showed AMR; graft survival rate decreased significantly in patients with AMR (59). However, protocol biopsies indicate that graft survival rates at 1 year were no different in patients with aTCMR than in those without rejection. This illustrates that evaluation of the rejection subtype is key, and that subclinical aTCMR and subclinical AMR should not be considered a single entity. Notably, all patients with subclinical aTCMR received steroid boluses according to routine practice (59). Favorable outcomes in patients with subclinical aTCMR could be explained by treatment effects, but no conclusions could be drawn about the impact of untreated subclinical aTCMR on graft outcome. Since borderline changes were not analyzed separately in this study, no conclusions can be drawn about the influence of subclinical borderline changes on outcome.

Although a randomized study showed that cyclosporine plus mycophenolate and steroids was associated with a higher risk for subclinical rejection (borderline changes and aTCMR) than tacrolimus, subclinical rejection was determined retrospectively, and was therefore untreated (60). Nevertheless, despite lack of treatment, this study showed that subclinical rejection did not lead to differences in graft functional evolution or fibrosis (60). Importantly, subclinical inflammation evaluated in non-fibrosed areas of protocol biopsies already displaying IFTA is more closely associated with poor graft survival than inflammation in otherwise normal biopsies (54, 58, 61–64). Finally, a randomized multicenter study from Canada indicated that subclinical TCMR or borderline changes occurred in 30–50% of patients at 6 months following transplant, depending on the level of tacrolimus dosing and the use of angiotensin-converting enzyme inhibitors/angiotensin II receptor 1 blockers (ACEi/ARBs), compared with other anti-hypertensive regimens. Patients with the lowest subclinical rejection risk (low-dose tacrolimus plus ACEi/ARBs) had reduced risk of progression of IFTA (65). It should be noted that this association did not necessarily indicate a causal relation between reduced risk of subclinical rejection and reduced IFTA progression.

Since the 2017 Banff Classification, caTCMR has been defined for the tubulointerstitial compartment, in part based on the grade of inflammation in atrophic areas (Banff i-IFTA score) (66). This histological phenotype in protocol biopsy studies is often preceded by interstitial inflammation in non-IFTA areas, coexists with interstitial inflammation in healthy areas, and constitutes a risk factor for progression of fibrosis and shortened graft survival (67, 68). Analysis of 1,500 1-year protocol biopsies by the Paris transplant group revealed that of the 893 biopsies scored as IFTA ≥1, 518 had no i-IFTA, 181 had an i-IFTA 1, and 194 had moderate to severe i-IFTA (2 or 3). Moderate to severe i-IFTA was associated with a decreased long-term graft survival rate and i-IFTA was superior to i, ti, and t scores for predicting allograft failure in patients with fibrosis at 1 year. In the Paris study, determinants of i-IFTA at 1 year were previous episodes of TCMR or BK virus nephropathy, as well as under-immunosuppression (67). As this indicates that both over- (e.g., polyomavirus nephropathy) and under-immunosuppression cause the same phenotype of chronic tubulointerstitial injury, it is difficult to establish the causes of that chronic injury.

Nankivell et al. reviewed i-IFTA in 2,481 biopsies from 362 patients, which were mainly protocol biopsies (mean number of seven biopsies per patient). Sequential histology demonstrated that interstitial inflammation occurred before the appearance of i-IFTA and chronic fibrosis. The 1-year i-IFTA intensity correlated with the number of prior TCMR episodes (68). In this study, i-IFTA was also associated with a worse graft survival rate and worse kidney function. Of note, however, although these data illustrated associations between caTCMR and graft failure, they were from retrospective studies. This limits the interpretability of results regarding whether it is necessary to treat, or not treat, such episodes of rejection. More recent data even indicate that most i-IFTA lesions are not preceded by rejection, and that even when they are, this rejection could be either TCMR or AMR (69). Finally, grade II caTCMR (chronic allograft arteriopathy, arterial intimal fibrosis with mononuclear cell inflammation in fibrosis, and formation of neointima) is even less well defined than other rejection subtypes, and may also be a manifestation of caAMR, cAMR, or mixed AMR/TCMR. Taken together, it is anticipated that further refinement of the diagnostic criteria of caTCMR will be important, both for clinical decision-making and before such criteria could be considered for clinical trial endpoints (70).

## INCOMPLETE INFLAMMATORY PHENOTYPES

Since tacrolimus and mycophenolate were introduced, the prevalence of tubulointerstitial inflammation in the first 2 years (subclinical aTCMR and borderline changes) has fallen from >50% to ~10% of transplant recipients (71). In addition, severity of inflammation has decreased to the point that subclinical aTCMR in protocol biopsies constitutes an uncommon diagnosis (63). Transplant biopsies with inflammation typically show changes or incomplete inflammatory phenotypes that are below the threshold for defining borderline changes (61, 66), raising the question whether such findings have any association with graft survival.

In an investigation of the clinical and pathological significance of borderline changes, lesions under i1 were excluded; when the significance of isolated tubulitis ($i = 0$, $t \geq 1$) on outcome was evaluated, no relationship was found between this lesion type and kidney allograft survival rate (51). Consequently, it was suggested not to include isolated tubulitis in the borderline category. This decision was agreed at Banff 2019, and included in the Banff criteria accordingly (72).

In another study including 200 of 275 patients with a 3-month protocol biopsy who did not meet the Banff criteria for TCMR grade IA, patients were classified as either no inflammation (i0t0) or inflammation ($i + t > 1$). Compared with transplant recipients without inflammation, those with inflammation showed higher chronic scores at 1 year, higher serum creatinine levels at 2 years, and higher incidence of *dn*DSA (73). In a further study, these authors illustrated that although these incomplete phenotypes of rejection were associated with increased risk of subsequent aTCMR, there was no association with worse graft survival (74). Notably, the lack of unified treatment protocol and small sample size hamper the interpretation of these results, and further research is warranted.

## IMPACT OF TREATING SUBCLINICAL INFLAMMATION

From the findings discussed above, we can conclude that the different histological phenotypes of subclinical inflammation in protocol biopsies – aTCMR, borderline changes, caTCMR, and interstitial inflammation without tubulitis—have been associated with decreased graft survival rates in retrospective and

observational cohort studies. Whether such subclinical inflammation should be recognized as pathology requiring treatment merits further discussion. A pioneering evaluation of steroid bolus treatment for kidney transplant recipients with subclinical aTCMR and borderline changes (t1/2/3 + i0/1 or t1 + i2/3) randomized participants to receive either biopsy at 1, 2, and 3 months with steroid treatment of subclinical inflammation, or no biopsy and no steroid treatment at 1, 2, and 3 months (75). Both groups had a 6-month protocol biopsy, received the then-standard of care immunosuppression with cyclosporine and azathioprine, and were followed for 2 years. Patients in the biopsy group had fewer cases of fibrosis at the 6-month protocol biopsy and better kidney function at 2 years, suggesting that treatment of subclinical inflammation preserves kidney structure and function. In this study, subclinical inflammation was present in ~50% of patients. These older data, with an outdated immunosuppressive regimen, suggested that detection of subclinical inflammation permits early, successful treatment and would be useful to include in future interventional trials. Another important weakness of this study is that it does not address the threshold of inflammation above or below which treatment improved (or failed to improve) kidney function at 2 years.

Subsequently, Kurtkoti et al. (76) designed a prospective randomized trial to evaluate whether treatment of rejection in protocol biopsies at 1 and 3 months preserved 1-year kidney function; participants also received cyclosporine and azathioprine. Rates of subclinical aTCMR and borderline changes were ~15% for each diagnosis at 1 and 3 months. The group of patients in whom treatment was adapted in response to protocol biopsy findings had better kidney function at 1 year, again suggesting that treating subclinical inflammation may improve outcome, although the effect on long-term graft failure was not studied.

As studies with older immunosuppressive regimens suggested treating subclinical inflammation with steroids (75, 76), Rush et al. performed a trial following the same design, but with a tacrolimus and mycophenolate immunosuppressive regimen (77). Although treating subclinical rejection (and rarely borderline changes) at 1, 2, or 3 months had no effect on interstitial fibrosis at 6 months or on kidney function at 12 and 24 months, the prevalence of subclinical inflammation at 1, 2, and 3 months was <10%, which was lower than expected. There is no mention of borderline changes that were not treated. Despite the low numbers of rejections, and in contrast to the hypothesis, the treatment arm tended to have higher chronic scores than the control arm, suggesting that treating subclinical rejection does not halt progressive chronic injury in patients receiving baseline immunosuppression with tacrolimus and mycophenolate (77). From this study, nothing can be implied about the impact of borderline changes in protocol biopsies early after transplantation.

Utility of 3- and 6-month protocol biopsies to predict graft survival was analyzed retrospectively in a pediatric population. Immunosuppression was increased (sometimes with steroid boluses) in patients with subclinical rejection. However, in those with borderline changes, treatment was selected by the attending physician; one-third of borderline episodes were not treated. The probability of reaching the composite outcome variable (i.e., an episode of clinical BPAR or graft failure in the next 5 years) was significantly higher in patients with untreated borderline changes than in treated patients (78). The retrospective nature of this study again warrants cautious interpretation of the data, especially regarding the effect of therapy, which was confounded by the decision of the attending physician.

Although it has been hypothesized that i-IFTA at 1 year is associated with under-immunosuppression (66), it is unclear whether increasing the immunosuppressive regimen or giving steroid-based antirejection therapy prevents or treats this condition. Importantly, immunosuppressive treatment may cause over-immunosuppression, which can create the same histological picture, through events such as the development of polyomavirus nephropathy.

In 1-year protocol biopsy studies performed in patients receiving current standard-of-care immunosuppression, the prevalence rates for subclinical inflammation are: ~3% for aTCMR (very low), 10%–15% for borderline changes, 15%–20% for incomplete phenotypes and 10% for caTCMR (51, 79, 80). The low frequency of aTCMR should be considered in studies that aim to reduce the incidence of TCMR further. Although treatment of aTCMR or borderline changes in protocol biopsies were suggested to be associated with improved outcome in an earlier era, this cannot be confirmed in studies using current immunosuppressive regimens. Previous heterogeneity in definitions of histological thresholds for the diagnosis of borderline changes (81) and interobserver variability also create additional problems for interstudy comparison.

Because of this heterogeneity in the literature, large variability in clinical practice remains (33–36). Some transplantation centers perform protocol biopsies and routinely treat subclinical aTCMR or borderline changes by increasing immunosuppression or using steroid boluses. Other centers would not treat borderline changes found on protocol biopsies unless there was additional evidence of rejection. A third group of centers do not perform protocol biopsies at all, and therefore never detect or treat subclinical changes. No data are available on European heterogeneity in this respect. Although i1t1 borderline changes in protocol biopsies are associated with a significant risk of subsequent TCMR, it remains unclear whether routinely treating caTCMR or incomplete phenotypes would improve graft outcome, as there is very limited literature on this phenotype.

## CONCLUSIONS

This paper reviews the history and classifications of TCMR and characterizes its potential role in clinical trials. ESOT has come to the following recommendations:

- BPAR and aTCMR are not equivalent: although many pivotal studies utilize BPAR to describe findings that could be aTCMR, some may be AMR or chronic rejection subtypes.
  - However, the specific definition of aTCMR has changed little over time, and can still broadly be used for longitudinal between-study comparisons.
- Acute TCMR (IA, IB, IIA, IIB, III) diagnosed in indication biopsies should remain included as a primary (non-inferiority) endpoint in clinical trials of kidney transplantation.

- Acute TCMR (IA, IB, IIA, IIB, III) diagnosed in protocol biopsies may be considered as a primary efficacy variable in clinical trials of kidney transplantation.
- Borderline changes (Banff Category 3 in the 2019 definition, restricted to Banff t ≥ 1 + i ≥ 1) diagnosed in indication biopsies following kidney transplantation are usually treated with antirejection therapy, and could be included as a primary (non-inferiority) efficacy variable in clinical trials for kidney transplantation.
  - Such borderline changes in indication biopsies are clinically relevant because of the impact of antirejection (immunosuppressive) therapy, which is required in a substantial number of cases diagnosed on indication biopsies.
- Diagnosis of at least one clinical episode of aTCMR or borderline changes (Banff Category 3 in the 2019 definition, restricted to Banff t ≥ 1 + i ≥ 1) in an indication biopsy, or aTCMR in a protocol biopsy, could be proposed as part of a composite primary efficacy endpoint in clinical trials aimed at preventing kidney transplant rejection.
- Few centers treat borderline changes identified in protocol biopsies with antirejection therapy; such changes are less clearly associated with outcome and should not be considered as primary efficacy measures in clinical trials for kidney transplantation.
- Awaiting further evidence, caTCMR (IA, IB, II) and tubulointerstitial inflammation below the Banff threshold for borderline changes should not be considered as measures of efficacy in clinical trials.

### Scientific Advice From the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA) with regard to These Conclusions

- The CHMP agreed with ESOT that the histological type of rejection (aTCMR, borderline changes, AMR) is a useful specification and that this detailing might be very informative in profiling efficacy of immunosuppression for kidney transplantation.
- The CHMP acknowledged the proposed clinically meaningful definition of borderline changes (to borderline suspicious for TCMR [restricted to Banff t ≥ 1 + i ≥ 1]) in indication biopsies.
  - However, in agreement with ESOT, the CHMP noted that there are clear between-center differences in performing protocol biopsies.
  - For regulatory purposes, the categorization of indication for renal transplant biopsy based on “per protocol” vs. “indication” may not be ideal.
- The CHMP commented that aTCMR (for both types of biopsies, protocol, and indication) and borderline changes (for indication biopsies only) could be primary efficacy endpoints for non-inferiority purposes, as the incidence of aTCMR is as low as 10%.
  - However, the concept of inferiority versus superiority is more applicable to the comparator type (approved vs standard of care) and not to the endpoint as such. Acceptance of a non-inferiority approach should be discussed *a priori*.
- The CHMP agreed that there is a need for more detailed analysis of the clinical relevance of minimal changes in the proposed histological subtypes of TCMR, including borderline changes in protocol biopsies.

## AUTHOR CONTRIBUTIONS

This is one of a series of papers written from the Broad Scientific Advice request submitted to the EMA/CHMP by ESOT: interactions between the EMA and ESOT regarding this request began in 2016. For the present article, working groups on histological and functional endpoints in kidney transplantation developed ESOT's position on the question “Does CHMP agree with the updated definitions of rejection and their potential use as primary endpoints in studies of kidney transplantation?”. The Centre for Evidence in Transplantation provided support with data extraction requests: these literature searches formed the basis of evidence used in the advice request and the article. Input into these outputs was provided from all ESOT members involved in the advice request process. The present article was adapted by MN from the final Broad Scientific Advice request submission (June 2020), presentation documents and minutes of the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020). The first draft of the article was further developed by DS and MR, then circulated to all authors for email review. The article was finalized and approved by all co-authors before submission for publication.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

MR has received lecture fees from Astellas and Chiesi; and research grant support (paid to institution) for investigator-initiated studies from Astellas and Chiesi. JB consults for Sanofi. GB has received honoraria and/or research funding from Astellas, CareDx, CSL Behring, Fresenius, Hansa, Neovii, and Vitaeris. KB has received honoraria and/or research funding from Alexion, Astellas, Bristol Myers Squibb, Chiesi, Fresenius, Hansa, Hexal, Merck, Novartis, Otsuka, Pfizer, Roche, Sandoz, Siemens, and Veloxis. DG reports consultancy agreements and scientific boards for AstraZeneca, BMS, Hansa, and Sanofi. LH reports speaker fees from Astellas, consultancy and research support from Chiesi, consultancy for Novartis, and research support from

Sandoz. RO has received grants/research support from Amgen, Astellas, and Chiesi; and speakers' bureaux/honoraria from Amgen, Astellas, Chiesi, Hansa, Neovii, Novartis, and Teva. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. European Medicines Agency. Clinical Investigation of Immunosuppressants for Solid Organ Transplantation (2008). CHMP/EWP/263148/06. Available from: https://www.ema.europa.eu/en/clinical-investigation-immunosuppressants-solid-organ-transplantation (Accessed October 18, 2021).
2. Meier-Kriesche H-U, Ojo AO, Hanson JA, Cibrik DM, Punch JD, Leichtman AB, et al. Increased Impact of Acute Rejection on Chronic Allograft Failure in Recent Era. *Transplantation* (2000) 70:1098–100. doi:10.1097/00007890-200010150-00018
3. Serón D, Arias M, Campistol JM, Morales JM. Late Renal Allograft Failure between 1990 and 1998 in Spain: a Changing Scenario. *Transplantation* (2003) 76:1588–94. doi:10.1097/01.TP.0000092495.07385.3C
4. Racusen LC, Colvin RB, Solez K, Mihatsch MJ, Halloran PF, Campbell PM, et al. Antibody-Mediated Rejection Criteria - an Addition to the Banff '97 Classification of Renal Allograft Rejection. *Am J Transpl* (2003) 3:708–14. doi:10.1034/j.1600-6143.2003.00072.x
5. Becker JU, Seron D, Rabant M, Roufosse C, Naesens M. Evolution of the Definition of Rejection in Kidney Transplantation and Its Use as an Endpoint in Clinical Trials. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10141
6. Mayrdorfer M, Liefeldt L, Wu K, Rudolph B, Zhang Q, Friedersdorff F, et al. Exploring the Complexity of Death-Censored Kidney Allograft Failure. *J Am Soc Nephrol* (2021) 32:1513–26. doi:10.1681/asn.2020081215
7. Mengel M, Sis B, Halloran PF. SWOT Analysis of Banff: Strengths, Weaknesses, Opportunities and Threats of the International Banff Consensus Process and Classification System for Renal Allograft Pathology. *Am J Transpl* (2007) 7:2221–6. doi:10.1111/j.1600-6143.2007.01924.x
8. Halloran P, Mathew T, Tomlanovich S, Groth C, Hooftman L, Barker C, et al. Mycophenolate Mofetil in Renal Allograft Recipients. *Transplantation* (1997) 63:39–47. doi:10.1097/00007890-199701150-00008
9. Pirsch JD, Miller J, Deierhoi MH, Vincenti F, Filo RS, et al. A Comparison of Tacrolimus (Fk506) and Cyclosporine for Immunosuppression after Cadaveric Renal Transplantation1. *Transplantation* (1997) 63:977–83. doi:10.1097/00007890-199704150-00013
10. Kahan BD. Efficacy of Sirolimus Compared with Azathioprine for Reduction of Acute Renal Allograft Rejection: a Randomised Multicentre Study. *The Lancet* (2000) 356:194–202. doi:10.1016/s0140-6736(00)02480-6
11. Naesens M, Thaunat O. BENEFIT of Belatacept: Kidney Transplantation Moves Forward. *Nat Rev Nephrol* (2016) 12:261–2. doi:10.1038/nrneph.2016.34
12. Ekberg H, Tedesco-Silva H, Demirbas A, Vítko Š, Nashan B, Gürkan A, et al. Reduced Exposure to Calcineurin Inhibitors in Renal Transplantation. *N Engl J Med* (2007) 357:2562–75. doi:10.1056/nejmoa067411
13. Budde K, Lehner F, Sommerer C, Reinke P, Arns W, Eisenberger U, et al. Five-Year Outcomes in Kidney Transplant Patients Converted from Cyclosporine to Everolimus: the Randomized ZEUS Study. *Am J Transpl* (2014) 15:119–28. doi:10.1111/ajt.12952
14. Rostaing L, Bunnapradist S, Grinyó JM, Ciechanowski K, Denny JE, Silva HT, Jr, et al. Novel Once-Daily Extended-Release Tacrolimus versus Twice-Daily Tacrolimus in De Novo Kidney Transplant Recipients: Two-Year Results of Phase 3, Double-Blind, Randomized Trial. *Am J Kidney Dis* (2016) 67:648–59. doi:10.1053/j.ajkd.2015.10.024
15. Silva HT, Yang HC, Meier-Kriesche H-U, Croy R, Holman J, Fitzsimmons WE, et al. Long-Term Follow-Up of a Phase III Clinical Trial Comparing Tacrolimus extended-release/MMF, tacrolimus/MMF, and Cyclosporine/MMF in De Novo Kidney Transplant Recipients. *Transplantation* (2014) 97:636–41. doi:10.1097/01.tp.0000437669.93963.8e
16. Krämer BK, Charpentier B, Bäckman L, Silva HT, Jr., Mondragon-Ramirez G, Cassuto-Viguier E, et al. Tacrolimus once Daily (ADVAGRAF) versus Twice Daily (PROGRAF) in De Novo Renal Transplantation: A Randomized Phase III Study. *Am J Transpl* (2010) 10:2632–43. doi:10.1111/j.1600-6143.2010.03256.x
17. Qazi Y, Shaffer D, Kaplan B, Kim DY, Luan FL, Peddi VR, et al. Efficacy and Safety of Everolimus Plus Low-Dose Tacrolimus versus Mycophenolate Mofetil Plus Standard-Dose Tacrolimus in De Novo Renal Transplant Recipients: 12-Month Data. *Am J Transpl* (2017) 17:1358–69. doi:10.1111/ajt.14090
18. McDonald S, Russ G, Campbell S, Chadban S. Kidney Transplant Rejection in Australia and New Zealand: Relationships between Rejection and Graft Outcome. *Am J Transpl* (2007) 7:1201–8. doi:10.1111/j.1600-6143.2007.01759.x
19. Pascual J, Berger SP, Witzke O, Tedesco H, Mulgaonkar S, Qazi Y, et al. Everolimus with Reduced Calcineurin Inhibitor Exposure in Renal Transplantation. *J Am Soc Nephrol* (2018) 29:1979–91. doi:10.1681/asn.2018010009
20. Arns W, Huppertz A, Rath T, Ziefle S, Rump LC, Hansen A, et al. Pharmacokinetics and Clinical Outcomes of Generic Tacrolimus (Hexal) versus Branded Tacrolimus in De Novo Kidney Transplant Patients. *Transplantation* (2017) 101:2780–8. doi:10.1097/tp.0000000000001843
21. Budde K, Bunnapradist S, Grinyo JM, Ciechanowski K, Denny JE, Silva HT, et al. Novel Once-Daily Extended-Release Tacrolimus (LCPT) versus Twice-Daily Tacrolimus in De Novo Kidney Transplants: One-Year Results of Phase III, Double-Blind, Randomized Trial. *Am J Transplant* (2014) 14:2796–806. doi:10.1111/ajt.12955
22. Halloran PF, Pereira AB, Chang J, Matas A, Picton M, De Freitas D, et al. Potential Impact of Microarray Diagnosis of T Cell-Mediated Rejection in Kidney Transplants: the INTERCOM Study. *Am J Transpl* (2013) 13:2352–63. doi:10.1111/ajt.12387
23. Mathew TH. A Blinded, Long-Term, Randomized Multicenter Study of Mycophenolate Mofetil in Cadaveric Renal Transplantation. *Transplantation* (1998) 65:1450–4. doi:10.1097/00007890-199806150-00007
24. Clayton PA, McDonald SP, Chapman JR, Chadban SJ. Mycophenolate versus Azathioprine for Kidney Transplantation. *Transplantation* (2012) 94:152–8. doi:10.1097/tp.0b013e31825475a3
25. Gaber AO, Moore LW, Alloway RR, Woodle ES, Pirsch J, Shihab F, et al. Acute Rejection Characteristics from a Prospective, Randomized, Double-Blind, Placebo-Controlled Multicenter Trial of Early Corticosteroid Withdrawal. *Transplantation* (2013) 95:573–9. doi:10.1097/tp.0b013e3182777efb
26. Clayton PA, McDonald SP, Russ GR, Chadban SJ. Long-term Outcomes after Acute Rejection in Kidney Transplant Recipients: an ANZDATA Analysis. *J Am Soc Nephrol* (2019) 30:1697–707. doi:10.1681/asn.2018111101
27. Vincenti F, Charpentier B, Vanrenterghem Y, Rostaing L, Bresnahan B, Darji P, et al. A Phase III Study of Belatacept-Based Immunosuppression Regimens versus Cyclosporine in Renal Transplant Recipients (BENEFIT Study). *Am J Transpl* (2010) 10:535–46. doi:10.1111/j.1600-6143.2009.03005.x

28. Vincenti F, Rostaing L, Grinyo J, Rice K, Steinberg S, Gaite L, et al. Belatacept and Long-Term Outcomes in Kidney Transplantation. *N Engl J Med* (2016) 374:333–43. doi:10.1056/nejmoa1506027
29. Bouatou Y, Viglietti D, Pievani D, Louis K, Duong Van Huyen JP, Rabant M, et al. Response to Treatment and Long-Term Outcomes in Kidney Transplant Recipients with Acute T Cell-Mediated Rejection. *Am J Transpl* (2019) 19: 1972–88. doi:10.1111/ajt.15299
30. Gaber LW, Moore LW, Alloway RR, Flax SD, Shokouh-Amiri MH, Schroder T, et al. Correlation between Banff Classification, Acute Renal Rejection Scores and Reversal of Rejection. *Kidney Int* (1996) 49:481–7. doi:10.1038/ki.1996.68
31. Minervini MI, Torbenson M, Scantlebury V, Vivas C, Jordan M, Shapiro R, et al. Acute Renal Allograft Rejection with Severe Tubulitis (Banff 1997 Grade IB). *Am J Surg Pathol* (2000) 24:553–8. doi:10.1097/00000478-200004000-00009
32. De Freitas DG, Sellarés J, Mengel M, Chang J, Hidalgo LG, Famulski KS, et al. The Nature of Biopsies with "Borderline Rejection" and Prospects for Eliminating This Category. *Am J Transpl* (2012) 12:191–201. doi:10.1111/j.1600-6143.2011.03784.x
33. Nankivell BJ, P'Ng CH, Chapman JR. Does Tubulitis without Interstitial Inflammation Represent Borderline Acute T Cell Mediated Rejection? *Am J Transpl* (2019) 19:132–44. doi:10.1111/ajt.14888
34. McRae M, Bouchard-Boivin F, Béland S, Noël R, Côté I, Lapointe I, et al. Impact of the Current versus the Previous Diagnostic Threshold on the Outcome of Patients With Borderline Changes Suspicious for T Cell-Mediated Rejection Diagnosed on Indication Biopsies. *Transplantation* (2018) 102:2120–5. doi:10.1097/tp.0000000000002327
35. Leblanc J, Subrt P, Paré M, Hartell D, Sénécal L, Blydt-Hansen T, et al. Practice Patterns in the Treatment and Monitoring of Acute T Cell-Mediated Kidney Graft Rejection in Canada. *Can J Kidney Health Dis* (2018) 5: 2054358117753616. doi:10.1177/2054358117753616
36. Beimler J, Zeier M. Borderline Rejection after Renal Transplantation - to Treat or Not to Treat. *Clin Transpl* (2009) 23(Suppl. 21):19–25. doi:10.1111/j.1399-0012.2009.01105.x
37. Wu K, Budde K, Lu H, Schmidt D, Liefeldt L, Glander P, et al. The Severity of Acute Cellular Rejection Defined by Banff Classification Is Associated with Kidney Allograft Outcomes. *Transplantation* (2014) 97:1146–54. doi:10.1097/01.tp.0000441094.32217.05
38. Wiebe C, Rush DN, Gibson IW, Pochinco D, Birk PE, Goldberg A, et al. Evidence for the Alloimmune Basis and Prognostic Significance of Borderline T Cell-Mediated Rejection. *Am J Transpl* (2020) 20:2499–508. doi:10.1111/ajt.15860
39. Canadian Multicentre Transplant Study Group. A Randomized Clinical Trial of Cyclosporine in Cadaveric Renal Transplantation. *N Engl J Med* (1983) 309: 809–15. doi:10.1056/NEJM198310063091401
40. Sutherland DER, Fryd DS, Strand MH, Canafax DM, Ascher NL, Payne WD, et al. Results of the Minnesota Randomized Prospective Trial of Cyclosporine versus Azathioprine-Antilymphocyte Globulin for Immunosuppression in Renal Allograft Recipients. *Am J Kidney Dis* (1985) 5:318–27. doi:10.1016/s0272-6386(85)80161-x
41. European Mycophenolate Mofetil Cooperative Study Group. Placebo-controlled Study of Mycophenolate Mofetil Combined with Cyclosporin and Corticosteroids for Prevention of Acute Rejection. European Mycophenolate Mofetil Cooperative Study Group. *Lancet* (1995) 345: 1321–5. doi:10.1016/S0140-6736(95)92534-1
42. Tricontinental Mycophenolate Mofetil Renal Transplantation Study Group. A Blinded, Randomized Clinical Trial of Mycophenolate Mofetil for the Prevention of Acute Rejection in Cadaveric Renal Transplantation. The Tricontinental Mycophenolate Mofetil Renal Transplantation Study Group. *Transplantation* (1996) 61:1029–37.
43. Sollinger HW. Mycophenolate Mofetil for the Prevention of Acute Rejection in Primary Cadaveric Renal Allograft Recipients. *Transplantation* (1995) 60: 225–32. doi:10.1097/00007890-199508000-00003
44. Vincenti F, Kirkman R, Light S, Bumgardner G, Pescovitz M, Halloran P, et al. Interleukin-2-Receptor Blockade with Daclizumab to Prevent Acute Rejection in Renal Transplantation. *N Engl J Med* (1998) 338:161–5. doi:10.1056/nejm199801153380304
45. Nashan Br., Light S, Hardie IR, Lin A, Johnson JR. Reduction of Acute Renal Allograft Rejection by Daclizumab1. *Transplantation* (1999) 67:110–5. doi:10.1097/00007890-199901150-00019
46. Kahan BD, Rajagopalan PR, Hall M. Reduction of the Occurrence of Acute Cellular Rejection Among Renal Allograft Recipients Treated with Basiliximab, a Chimeric Anti-interleukin-2-receptor Monoclonal Antibody1,2. *Transplantation* (1999) 67:276–84. doi:10.1097/00007890-199901270-00016
47. Nashan B, Moore R, Amlot P, Schmidt A-G, Abeywickrama K, Soulillou J-P. Randomised Trial of Basiliximab versus Placebo for Control of Acute Cellular Rejection in Renal Allograft Recipients. *The Lancet* (1997) 350:1193–8. doi:10.1016/s0140-6736(97)09278-7
48. Durrbach A, Pestana JM, Pearson T, Vincenti F, Garcia VD, Campistol J, et al. A Phase III Study of Belatacept versus Cyclosporine in Kidney Transplants from Extended Criteria Donors (BENEFIT-EXT Study). *Am J Transpl* (2010) 10:547–57. doi:10.1111/j.1600-6143.2010.03016.x
49. Budde K, Becker T, Arns W, Sommerer C, Reinke P, Eisenberger U, et al. Everolimus-based, Calcineurin-inhibitor-free Regimen in Recipients of De-novo Kidney Transplants: an Open-Label, Randomised, Controlled Trial. *The Lancet* (2011) 377:837–47. doi:10.1016/s0140-6736(10)62318-5
50. Meehan SM, Siegel CT, Aronson AJ, Bartosh SM, Thistlethwaite JR, Woodle ES, et al. The Relationship of Untreated Borderline Infiltrates by the Banff Criteria to Acute Rejection in Renal Allograft Biopsies. *J Am Soc Nephrol* (1999) 10:1806–14. doi:10.1681/asn.v1081806
51. Nankivell BJ, Agrawal N, Sharma A, Taverniti A, P'Ng CH, Shingde M, et al. The Clinical and Pathological Significance of Borderline T Cell-Mediated Rejection. *Am J Transpl* (2018) 19:1452–63. doi:10.1111/ajt.15197
52. Shishido S, Asanuma H, Nakai H, Mori Y, Satoh H, Kamimaki I, et al. The Impact of Repeated Subclinical Acute Rejection on the Progression of Chronic Allograft Nephropathy. *J Am Soc Nephrol* (2003) 14:1046–52. doi:10.1097/01.asn.0000056189.02819.32
53. Alperovich G, Maldonado R, Moreso F, Fulladosa X, Grinyó JM, Serón D. Glomerular Enlargement Assessed by Paired Donor and Early Protocol Renal Allograft Biopsies. *Am J Transpl* (2004) 4:650–4. doi:10.1111/j.1600-6143.2004.00392.x
54. Azevedo F, Alperovich G, Moreso F, Ibernon M, Gomà M, Fulladosa X, et al. Glomerular Size in Early Protocol Biopsies Is Associated with Graft Outcome. *Am J Transpl* (2005) 5:2877–82. doi:10.1111/j.1600-6143.2005.01126.x
55. Wiebe C, Gibson IW, Blydt-Hansen TD, Karpinski M, Ho J, Storsley LJ, et al. Evolution and Clinical Pathologic Correlations of De Novo Donor-specific HLA Antibody post Kidney Transplant. *Am J Transpl* (2012) 12:1157–67. doi:10.1111/j.1600-6143.2012.04013.x
56. García-Carro C, Dörje C, Åsberg A, Midtvedt K, Scott H, Reinholt FP, et al. Inflammation in Early Kidney Allograft Surveillance Biopsies with and without Associated Tubulointerstitial Chronic Damage as a Predictor of Fibrosis Progression and Development of De Novo Donor Specific Antibodies. *Transplantation* (2017) 101:1410–5. doi:10.1097/tp.0000000000001216
57. Choi BS, Shin MJ, Shin SJ, Kim YS, Choi YJ, Kim Y-S, et al. Clinical Significance of an Early Protocol Biopsy in Living-Donor Renal Transplantation: Ten-Year Experience at a Single center. *Am J Transpl* (2005) 5:1354–60. doi:10.1111/j.1600-6143.2005.00830.x
58. Mengel M, Gwinner W, Schwarz A, Bajeski R, Franz I, Bröcker V, et al. Infiltrates in Protocol Biopsies from Renal Allografts. *Am J Transpl* (2007) 7: 356–65. doi:10.1111/j.1600-6143.2006.01635.x
59. Loupy A, Vernerey D, Tinel C, Aubert O, Duong van Huyen J-P, Rabant M, et al. Subclinical Rejection Phenotypes at 1 Year post-transplant and Outcome of Kidney Allografts. *J Am Soc Nephrol* (2015) 26:1721–31. doi:10.1681/asn.2014040399
60. Scholten EM, Rowshani AT, Cremers S, Bemelman FJ, Eikmans M, van Kan E, et al. Untreated Rejection in 6-month Protocol Biopsies Is Not Associated with Fibrosis in Serial Biopsies or with Loss of Graft Function. *J Am Soc Nephrol* (1200) 17:2622–32. doi:10.1681/ASN.2006030227
61. Gago M, Cornell LD, Kremers WK, Stegall MD, Cosio FG. Kidney Allograft Inflammation and Fibrosis, Causes and Consequences. *Am J Transpl* (2012) 12: 1199–207. doi:10.1111/j.1600-6143.2011.03911.x

62. Cosio FG, Grande JP, Wadei H, Larson TS, Griffin MD, Stegall MD. Predicting Subsequent Decline in Kidney Allograft Function from Early Surveillance Biopsies. *Am J Transpl* (2005) 5:2464–72. doi:10.1111/j.1600-6143.2005.01050.x
63. Moreso F, Ibernon M, Gomà M, Carrera M, Fulladosa X, Hueso M, et al. Subclinical Rejection Associated with Chronic Allograft Nephropathy in Protocol Biopsies as a Risk Factor for Late Graft Loss. *Am J Transpl* (2006) 6:747–52. doi:10.1111/j.1600-6143.2005.01230.x
64. Park WD, Griffin MD, Cornell LD, Cosio FG, Stegall MD. Fibrosis with Inflammation at One Year Predicts Transplant Functional Decline. *J Am Soc Nephrol* (2010) 21:1987–97. doi:10.1681/asn.2010010049
65. Cockfield SM, Wilson S, Campbell PM, Cantarovich M, Gangji A, Houde I, et al. Comparison of the Effects of Standard vs Low-dose Prolonged-release Tacrolimus with or without ACEi/ARB on the Histology and Function of Renal Allografts. *Am J Transpl* (2019) 19:1730–44. doi:10.1111/ajt.15225
66. Haas M, Loupy A, Lefaucheur C, Roufosse C, Glotz D, Seron D, et al. The Banff 2017 Kidney Meeting Report: Revised Diagnostic Criteria for Chronic Active T Cell-Mediated Rejection, Antibody-mediated Rejection, and Prospects for Integrative Endpoints for Next-generation Clinical Trials. *Am J Transpl* (2018) 18:293–307. doi:10.1111/ajt.14625
67. Lefaucheur C, Gosset C, Rabant M, Viglietti D, Verine J, Aubert O, et al. T Cell-Mediated Rejection Is a Major Determinant of Inflammation in Scarred Areas in Kidney Allografts. *Am J Transpl* (2018) 18:377–90. doi:10.1111/ajt.14565
68. Nankivell BJ, Shingde M, Keung KL, Fung CLS, Borrows RJ, O'Connell PJ, et al. The Causes, Significance and Consequences of Inflammatory Fibrosis in Kidney Transplantation: The Banff i-IFTA Lesion. *Am J Transpl* (2018) 18:364–76. doi:10.1111/ajt.14609
69. Helgeson ES, Mannon R, Grande J, Gaston RS, Cecka MJ, Kasiske BL, et al. I-IFTA and Chronic Active T Cell-Mediated Rejection: A Tale of 2 (DeKAF) Cohorts. *Am J Transpl* (2021) 21:1866–77. doi:10.1111/ajt.16352
70. Naesens M, Haas M, Loupy A, Roufosse C, Mengel M. Does the Definition of Chronic Active T Cell-Mediated Rejection Need Revisiting? *Am J Transpl* (2021) 21:1689–90. doi:10.1111/ajt.16419
71. Torres IB, Reisaeter AV, Moreso F, Åsberg A, Vidal M, Garcia-Carro C, et al. Tacrolimus and Mycophenolate Regimen and Subclinical Tubulo-Interstitial Inflammation in Low Immunological Risk Renal Transplants. *Transpl Int* (2017) 30:1119–31. doi:10.1111/tri.13002
72. Loupy A, Haas M, Roufosse C, Naesens M, Adam B, Afrouzian M, et al. The Banff 2019 Kidney Meeting Report (I): Updates on and Clarification of Criteria for T Cell- and Antibody-mediated Rejection. *Am J Transpl* (2020) 20:2318–31. doi:10.1111/ajt.15898
73. Mehta R, Bhusal S, Randhawa P, Sood P, Cherukuri A, Wu C, et al. Short-term Adverse Effects of Early Subclinical Allograft Inflammation in Kidney Transplant Recipients with a Rapid Steroid Withdrawal Protocol. *Am J Transpl* (2018) 18:1710–7. doi:10.1111/ajt.14627
74. Mehta RB, Tandukar S, Jorgensen D, Randhawa P, Sood P, Puttarajappa C, et al. Early Subclinical Tubulitis and Interstitial Inflammation in Kidney Transplantation Have Adverse Clinical Implications. *Kidney Int* (2020) 98:436–47. doi:10.1016/j.kint.2020.03.028
75. Rush D, Nickerson P, Gough J, McKenna R, Grimm P, Cheang M, et al. Beneficial Effects of Treatment of Early Subclinical Rejection: a Randomized Study. *J Am Soc Nephrol* (1998) 9:2129–34. doi:10.1681/asn.v9112129
76. Kurtkoti J, Sakhuja V, Sud K, Minz M, Nada R, Kohli HS, et al. The Utility of 1- and 3-month Protocol Biopsies on Renal Allograft Function: a Randomized Controlled Study. *Am J Transpl* (2008) 8:317–23. doi:10.1111/j.1600-6143.2007.02049.x
77. Rush D, Arlen D, Boucher A, Busque S, Cockfield SM, Girardin C, et al. Lack of Benefit of Early Protocol Biopsies in Renal Transplant Patients Receiving TAC and MMF: a Randomized Study. *Am J Transpl* (2007) 7:2538–45. doi:10.1111/j.1600-6143.2007.01979.x
78. Seifert ME, Yanik MV, Feig DI, Hauptfeld-Dolejsek V, Mroczek-Musulman EC, Kelly DR, et al. Subclinical Inflammation Phenotypes and Long-Term Outcomes after Pediatric Kidney Transplantation. *Am J Transpl* (2018) 18:2189–99. doi:10.1111/ajt.14933
79. Moreso F, Serón D, Carrera M, Gil-Vernet S, Cruzado JM, Hueso M, et al. Baseline Immunosuppression Is Associated with Histological Findings in Early Protocol Biopsies. *Transplantation* (2004) 78:1064–8. doi:10.1097/01.tp.0000137268.85155.11
80. Ortiz F, Gelpi R, Helanterä I, Melilli E, Honkanen E, Bestard O, et al. Decreased Kidney Graft Survival in Low Immunological Risk Patients Showing Inflammation in normal Protocol Biopsies. *PLoS One* (2016) 11:e0159717. doi:10.1371/journal.pone.0159717
81. Becker JU, Chang A, Nickeleit V, Randhawa P, Roufosse C. Banff Borderline Changes Suspicious for Acute T Cell-Mediated Rejection: where Do We Stand? *Am J Transpl* (2016) 16:2654–60. doi:10.1111/ajt.13784

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10140

# Proposed Definitions of Antibody-Mediated Rejection for Use as a Clinical Trial Endpoint in Kidney Transplantation

*Candice Roufosse[1†], Jan Ulrich Becker[2†], Marion Rabant[3], Daniel Seron[4], Maria Irene Bellini[5], Georg A. Böhmig[6], Klemens Budde[7], Fritz Diekmann[8], Denis Glotz[9], Luuk Hilbrands[10], Alexandre Loupy[11], Rainer Oberbauer[6], Liset Pengel[12], Stefan Schneeberger[13] and Maarten Naesens[14]**

[1]Department of Immunology and Inflammation, Centre for Inflammatory Disease, Imperial College London, London, United Kingdom, [2]Institute of Pathology, University Hospital Cologne, Cologne, Germany, [3]Department of Pathology, Hôpital Necker-Enfants Malades, Paris, France, [4]Department of Nephrology and Kidney Transplantation, Vall d'Hebrón University Hospital, Barcelona, Spain, [5]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy, [6]Division of Nephrology and Dialysis, Department of Internal Medicine, Medical University of Vienna, Vienna, Austria, [7]Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany, [8]Department of Nephrology and Kidney Transplantation, Hospital Clinic Barcelona, Barcelona, Spain, [9]Paris Translational Research Center for Organ Transplantation, Hôpital Saint Louis, Paris, France, [10]Department of Nephrology, Radboud University Medical Center, Nijmegen, Netherlands, [11]Paris Translational Research Center for Organ Transplantation, Hôpital Necker, Paris, France, [12]Centre for Evidence in Transplantation, University of Oxford, Oxford, United Kingdom, [13]Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria, [14]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

Antibody-mediated rejection (AMR) is caused by antibodies that recognize donor human leukocyte antigen (HLA) or other targets. As knowledge of AMR pathophysiology has increased, a combination of factors is necessary to confirm the diagnosis and phenotype. However, frequent modifications to the AMR definition have made it difficult to compare data and evaluate associations between AMR and graft outcome. The present paper was developed following a Broad Scientific Advice request from the European Society for Organ Transplantation (ESOT) to the European Medicines Agency (EMA), which explored whether updating guidelines on clinical trial endpoints would encourage innovations in kidney transplantation research. ESOT considers that an AMR diagnosis must be based on a combination of histopathological factors and presence of donor-specific HLA antibodies in the recipient. Evidence for associations between individual features of AMR and impaired graft outcome is noted for microvascular inflammation scores ≥2 and glomerular basement membrane splitting of >10% of the entire tuft in the most severely affected glomerulus. Together, these should form the basis for AMR-related endpoints in clinical trials of kidney transplantation, although modifications and restrictions to the Banff diagnostic definition of AMR are proposed for this purpose. The EMA provided recommendations based on this Broad Scientific Advice request in December 2020; further discussion, and consensus on the restricted definition of the AMR endpoint, is required.

**Keywords: kidney transplantation, outcomes, biopsy, histology, antibody-mediated rejection, EMA guideline**

OPEN ACCESS

***Correspondence:**
Maarten Naesens
maarten.naesens@kuleuven.be

[†]These authors have contributed equally to this work

**Received:** 21 October 2021
**Accepted:** 03 March 2022
**Published:** 20 May 2022

**Citation:**
Roufosse C, Becker JU, Rabant M, Seron D, Bellini MI, Böhmig GA, Budde K, Diekmann F, Glotz D, Hilbrands L, Loupy A, Oberbauer R, Pengel L, Schneeberger S and Naesens M (2022) Proposed Definitions of Antibody-Mediated Rejection for Use as a Clinical Trial Endpoint in Kidney Transplantation. Transpl Int 35:10140. doi: 10.3389/ti.2022.10140

## WHAT IS ANTIBODY-MEDIATED REJECTION?

Although biopsy-proven acute rejection (BPAR) remains widely used as a primary efficacy variable in the clinical trial setting (1), it is a non-specific term. Despite often considered equivalent to acute T cell-mediated rejection (aTCMR), BPAR likely also includes unrecognized cases of antibody-mediated injury, especially in research published in the twentieth century. Antibody-mediated rejection (AMR), distinct from hyperacute rejection, emerged as a diagnostic concept in 1997 (2); subsequently it was recognized as a frequent cause of graft failure and an important cause of post-transplant complications (3–7). Affecting up to 25% of kidney allograft recipients (8, 9), the risk for AMR is low in the first year post transplantation in pre-transplant donor-specific antibody (DSA)-negative patients but reaches 30–40% in those who are DSA+. Beyond the first year following transplantation, risk for developing *de novo* (*dn*)DSA and subsequent AMR is associated with insufficient immunosuppression, which can result—among other factors—from non-adherence to standard-of-care regimens (10).

Advances in the development of sensitive assays for DSA identification have improved our understanding of AMR histopathology (11, 12). AMR is caused by antibodies that recognize donor human leukocyte antigen (HLA) on the kidney allograft endothelium, foreign to the recipient. Antibodies can also be formed against other allogeneic targets including non-HLA antibodies (e.g., against minor histocompatibility antigens) or non-allogeneic targets such as endothelial antigens or vimentin (13). DSA may develop before transplantation (because of blood transfusion, pregnancy, or previous allografts), or afterwards (*dn*DSA). AMR is recognized as a spectrum of discrete injury patterns, as outlined below.

## AMR IN THE BANFF CLASSIFICATION

The detrimental impact of AMR on kidney transplantation outcome has been known for decades, as illustrated by the early routine implementation of crossmatching to avoid this rejection phenotype (14). The theoretical importance of AMR in kidney transplantation pathology was acknowledged at the first Banff meeting to focus on allograft pathology, in 1991 (15), However, this report only designated hyperacute rejection because of preformed DSA as a separate category (category 2) that was recognized as the most severe form of rejection, usually leading to immediate graft loss (15). In addition to hyperacute rejection, the 1997 update included delayed (accelerated acute) AMR and described histopathological and serological (crossmatch) diagnostic criteria (2). Reflecting the growing body of knowledge about AMR in kidney transplantation, diagnostic criteria and subcategories of AMR in Banff Classifications have changed considerably over time.

The next advancement followed the introduction of C4d staining, which documented histopathogenetic links between circulating DSA and organ damage, by detecting complement activation by DSA fixed to surface antigens on the endothelial cell (16). The 2001 Banff meeting recognized several histological types of acute/active (a)AMR, thereby expanding Category 2 diagnoses to include the following: 1) acute tubular necrosis-like minimal inflammation; 2) with capillary margination (glomerulitis and peritubular capillaritis [now considered microvascular inflammation, MVI]) and/or thromboses; and 3) with transmural arteritis and/or arterial wall necrosis. The reference to clinical presentations ("hyperacute" and "accelerated acute") was abandoned, with emphasis shifting to histopathological features. Of note, all three AMR subtypes required C4d positivity (17).

Chronic (c)AMR subtypes were first recognized in the Banff 2005 update, as chronic active (ca)AMR, with transplant glomerulopathy (TG) and/or severe peritubular capillary basement membrane multilayering (PTCML) and/or simple interstitial fibrosis and tubular atrophy and/or arterial fibrous intimal thickening, with C4d positivity (18). Evidence of the pathogenetic link between aAMR and cAMR was discussed at the Banff 2017 meeting (19). The requirement for both DSA and C4d positivity to diagnose all subcategories of AMR (18) was relaxed in 2009, when subcategories for C4d−/DSA+ cases "suspicious for AMR" were created, matching the morphological patterns listed above but without C4d positivity (20). Further evidence (4) led to full recognition of C4d− AMR in the 2013 Banff update (21), and a diagnostic flowchart was created featuring subcategories "C4d positivity without evidence of rejection," "suspicious for aAMR," "aAMR," "suspicious for caAMR," "caAMR," and "cAMR." The flowchart accommodated numerous combinations of histopathological findings (21) that were simplified in the 2017 Banff update to form the categories listed in **Table 1** (19, 22, 23). Most importantly, the "suspicious" categories were abandoned. Subsequently, only minor modifications have been made (22).

Currently, AMR diagnosis within Banff Classification category 2 is based on four partially overlapping components: histological features of AMR activity; evidence of antibody interaction with graft vascular endothelium; histological features of AMR chronicity; and DSA or equivalents (**Table 1**) (19, 21). Reaching an AMR diagnosis requires a combination of these criteria to be fulfilled.

## AMR AND ALLOGRAFT OUTCOME

Updates to the Banff Classification of AMR over time make it difficult to maintain long-term follow-up registries or compare literature. Moreover, the interobserver agreement (κ-statistic) of the most important lesion scores for AMR was quoted as 0.39 for Banff Lesion Score g, 0.38 for ptc, and 0.48 for cg—at best a fair-to-moderate agreement, even among very experienced transplant nephropathologists (24).

Several problems arise when reviewing evidence of an association between AMR and allograft outcome. Firstly, AMR definitions have changed very frequently since 2001, as outlined above, making it difficult to compare data from studies conducted over the last 2 decades. Secondly, Banff diagnostic criteria and categories are adjusted based on antecedent literature, and as they arise as a synthesis of several different studies, rarely fully align

**TABLE 1 |** Antibody-mediated changes (19, 22, 23); diagnostic criteria groups are used to reach one diagnosis.

| Diagnosis | Diagnostic criteria groups |
|---|---|
| **C4d staining without evidence of rejection**<br>• Lesion Score C4d > 1 (immunofluorescence on fresh frozen tissue) OR C4d > 0 (immunohistochemistry on paraffin-embedded tissue) AND<br>• Banff Lesion Scores t0, v0, no arterial intimal fibrosis with mononuclear cell inflammation in fibrosis and formation of neointima, no criterion from Group 1 (AMR Banff activity), no criterion from Group 4 (histologic features of AMR chronicity), no increased expression of thoroughly validated gene transcripts/classifiers in the biopsy tissue strongly associated with AMR | **Group 1: AMR activity**<br>• Banff Lesion Score g > 0 in the absence of glomerulonephritis and/or Banff Lesion Score ptc>0 in the absence of TCMR or borderline changes<br>• Banff Lesion Score v > 0<br>• Acute thrombotic microangiopathy in the absence of any other cause<br>• Acute tubular injury in the absence of any other apparent cause |
| **Active AMR**<br>• No criterion from Group 4 (histologic features of AMR chronicity) AND<br>• ≥1 criterion from Group 1 (AMR activity) AND<br>• ≥1 criterion from Group 2 (antibody interaction with tissue) AND<br>• ≥1 criterion from Criteria Group 3 (DSA or equivalent) AND | **Group 2: Antibody interaction with tissue Banff Lesion Score C4d > 1 (IF on fresh frozen tissue) or C4d > 0 (IHC on paraffin-embedded tissue)**<br>• At least moderate microvascular inflammation (g + ptc>1) in the absence of borderline changes (Diagnostic Category 3) or acute TCMR (aTCMR; Diagnostic Category 4). If borderline changes or aTCMR are present, Banff Lesion Score g + ptc>1 is not sufficient; g ≥ 1 is required<br>• Increased expression of thoroughly validated gene transcripts/classifiers in the biopsy tissue strongly associated with AMR |
| **Chronic active AMR**<br>• ≥1 feature from Group 4 (histologic features of AMR chronicity) AND<br>• ≥1 criterion from Group 2 (antibody interaction with tissue) AND<br>• ≥1 criterion from Group 3 (DSA or equivalent) | **Group 3: DSA or equivalent**<br>• DSA (anti-HLA or other specificity)<br>• Banff Lesion Score C4d > 1 (IF on fresh frozen tissue) or C4d > 0 (IHC on paraffin-embedded tissue)<br>• Increased expression of thoroughly validated gene transcripts/classifiers in the biopsy tissue strongly associated with AMR |
| **Chronic AMR**<br>• Banff 2017 permits the use of this term for biopsy specimens showing TG and/or PTCML in the absence of criterion of current/recent antibody interaction with the endothelium (Criteria Group 2) but with a prior documented diagnosis of active or chronic active AMR or documented prior evidence of DSA | **Group 4: Histologic features of AMR chronicity**<br>• Banff Lesion Score cg > 0 (by light microscopy or EM, if available), excluding biopsies with evidence of chronic thrombotic microangiopathy<br>• ≥7 layers in 1 cortical peritubular capillary and ≥5 in 2 additional capillaries, avoiding portions cut tangentially by EM, if available (severe PTCML); arterial intimal fibrosis of new onset, excluding other causes; leukocytes within the sclerotic intima favor chronic AMR if there is no prior history of biopsy-proven TCMR with arterial involvement, but are not required |

*AMR, antibody-mediated rejection; aTCMR, acute T cell-mediated rejection; DSA, donor-specific antibody; EM, electron microscopy; HLA, human leukocyte antigen; IF, immunofluorescence; IHC, immunohistochemistry; PTCML, peritubular capillary basement membrane multilayering; TCMR, T cell-mediated rejection; TG, transplant glomerulopathy. Adapted from Roufosse et al., 2018 (23).*

with literature on which they are based. Conversely, outcomes of AMR diagnoses according to their strict definitions in the Banff Classification have rarely been investigated for either their association with outcome, or their success in delineating which patients are eligible for a specific therapy; where this has been done, results show improved prediction of outcome with the 2013 version compared to the 2003/2007 Classification, and with the 2013 version compared to the 2017 Classification (25, 26). Instead, researchers often use slightly different definitions for the categories, with bespoke combinations of components and cut-offs for defining AMR, instead of the strict definitions last proposed by Banff. Thirdly, several Banff inclusion criteria for defining AMR are difficult to apply in clinical practice; this is one reason why precise Banff categories for AMR are rarely tested for their association with outcome. For example, cAMR categorization is mainly based on light microscopic features of Banff Lesion Score cg (TG), because detection of "severe PTCML" as an inclusion criterion for cAMR relies on electron microscopy (EM) (27). Few studies use arterial intimal thickening of new onset as an inclusion criterion for cAMR because it is difficult to score: this finding is dependent on arteries being cut transversally, is associated with unreliable arterial sampling (of both the current and previous biopsies), and in some cases is impossible to obtain because of lack of previous biopsies to use as a baseline. "Acute TMA" (thrombotic microangiopathy) is rarely the sole inclusion criterion for AMR, because it is hard to completely exclude TMA of other causes, it is rarely seen as an isolated feature without other features of AMR such as microcirculation inflammation, and because a Banff Working Group has yet to agree on a consensus definition of TMA (68).

To our knowledge, no method of transcriptome analysis has been formally recognized as thoroughly validated by Banff. No transplant centers have obtained adequate clinical validation to use transcript analysis for defining AMR, as required by Banff consensus. In addition, although the Banff Classification makes no distinction between AMR in patients with preformed DSA (high-risk) and non-sensitized (low-risk) transplant recipients, the diagnosis and treatment pathways for both groups might be quite different, as are the underlying biology and clinical phenotypes. Consequently, AMR classification may need to include more than histology, because identical histological diagnoses in the kidney (such as TG or TMA) can be the consequence of different disease entities (28). Finally, a

diagnosis based on histology alone is not sufficient to describe the underlying pathophysiology. As suggested in the consensus report (28), for disease classification and outcome prediction, timing and clinical phenotype are crucial; and whether the patient has *dn*DSA, preformed DSA, or no HLA-DSA must also be known.

Thus, although the basic principles of diagnosing AMR have generally remained constant, given the considerable changes to Banff definitions of AMR, longitudinal comparison of literature findings is more challenging for AMR than for aTCMR. Interpretation of the AMR literature must be undertaken cautiously, taking account of these limitations. In reviewing evidence that could serve as background for the definition of AMR, first it is important to evaluate studies that have assessed outcomes associated with various combinations of biopsy features. In the following sections, we divide this information into subcategories broadly based on the Banff classification. After evaluating the literature on allograft outcome, we consider data relating to associations between outcome and individual biopsy features that are components of AMR (19). Moreover, it must be stated that we had to use the best available evidence for our consensus definitions of AMR. Inevitably, we had to omit rarer, insufficiently defined or researched phenotypes of the wide clinicopathological spectrum of AMR. Research should focus on diagnostic criteria for such rarer phenotypes, their outcome and their suitability for inclusion in AMR treatment studies. Of course, both the Banff Classification and future endpoint definitions will reflect any such evidence arising from these studies. In the interim, researchers are free to use their own endpoints. The choice of alternative endpoints is particularly justified in special scenarios such as in sensitized recipients requiring desensitization for transplantation.

## BANFF CLASSIFICATION: AMR SUBCATEGORIES AS ENDPOINTS

### C4d Staining Without Evidence of Rejection

This subcategory is discussed in conjunction with C4d positivity with acute tubular injury (ATI) in the absence of other apparent cause.

### Active AMR

Much of the evidence for an association between aAMR and outcome (i.e., graft loss) comes from publications that only include components of aAMR (e.g., MVI, C4d, and/or DSA) (**Table 2**) (29–34). Evidence of an association between aAMR and outcome derives from retrospective observational studies that rarely distinguish between pure aAMR and caAMR; thus, few studies specifically investigate the relationship between a Banff subcategory aAMR diagnosis and outcome.

Given the heterogeneity of definitions, overall, the quality of evidence that strictly defines the association between aAMR and increased graft loss is low; in recipients of HLA-incompatible grafts, quality of evidence is higher. However, if one evaluates the literature with less stringency about the exact AMR definition, there is consensus that aAMR is an important risk factor for graft failure (10, 28). Moreover, in the era of powerful T cell inhibition as standard immunosuppression, outcome after aAMR at time of graft dysfunction is significantly worse than outcome after aTCMR (35).

In the absence of dysfunction (i.e., subclinical AMR in protocol biopsies), the outcome used for features of aAMR is usually TG rather than graft loss, although a retrospective study indicated that subclinical AMR in 1-year protocol biopsies had a detrimental impact on graft survival (36). There is general agreement on treating aAMR regardless of whether graft dysfunction occurs, further illustrating the clinical relevance of this phenotype (28). This is discussed further in the section below, "Subclinical AMR Including Incomplete Phenotypes."

Our proposal is that aAMR, exactly as defined by the current Banff classification, cannot be adopted as a surrogate endpoint for future cAMR and graft loss in low-risk situations, i.e., non-sensitized graft recipients, without DSA against the graft. Conversely, in high-risk patients, i.e., sensitized patients with DSA against the graft, evidence supports features of aAMR as a surrogate for graft loss, especially if associated with graft dysfunction. Future research should aim to establish outcome (graft loss, graft function, future cAMR, or caAMR) in patients with aAMR, strictly defined according to Banff criteria and specifically excluding cases with features of chronicity. Such research should involve retrospective and prospective studies, and high- and conventional-risk transplantations. Data from randomized controlled trials investigating aAMR treatment regimens would also be particularly valuable. Although further data are awaited, there is broad consensus on the clinical relevance and impact of aAMR after kidney transplantation. Since aAMR leads to therapeutic interventions, treatment burden, associated morbidity, and increased cost, features of aAMR represent a key endpoint for interventional studies.

### Chronic AMR and Chronic Active AMR

According to the 2017 Banff Classification, a diagnosis of cAMR or caAMR can only be established based on presence of TG (Banff Lesion Score cg > 0) and/or severe PTCML. For caAMR, this must be accompanied by "evidence of antibody interaction with tissue" and "DSA or equivalent"; for cAMR, this must be in conjunction with "a prior documented diagnosis of aAMR or caAMR or documented prior evidence of DSA." The ill-defined transplant vasculopathy is no longer considered a chronicity parameter for these diagnoses (19). Data for patients fulfilling identical Banff criteria of severe PTCML as the indicator of AMR chronicity, in conjunction with solid-phase DSA testing, are scant. Therefore, we present evidence only for outcomes in patients with TG, with "histological lesions strongly associated with AMR" (i.e., MVI ≥2, C4d positivity, or "increased expression of thoroughly validated gene transcripts/classifiers in biopsy tissue") (**Table 3**) (37–54). In reviewing literature on cg, as with the other histological lesions, caution should be taken because of the relatively limited interobserver agreement (24).

A retrospective study of 44 patients with TG examined the outcome of graft loss, with TG defined as Banff cg > 0 (glomerular basement membrane splitting of >10% of the entire tuft in the most severely affected glomerulus) (2); this definition remained relevant until Banff 2011 (55). With this TG threshold—higher

**TABLE 2 |** Studies investigating associations between Banff diagnostic category "active AMR" and outcome (29–34).

| References | Endpoint | Definition of aAMR | Study type | Cohort | Findings | Level of evidence (grade) |
|---|---|---|---|---|---|---|
| Solar-Cafaggi et al. (29) | Graft loss | Mixed (unseparated) aAMR and caAMR (Banff 2007 or 2017 criteria); indication and protocol biopsies; incomplete DSA data | Single-center retrospective | N = 201 | Increased graft loss ($p = 0.001$) | Very low |
| Sai et al. (30) | sCr × 1.5; cAMR and graft loss | Banff 2013 | Retrospective | N = 627 protocol + indication biopsies; C4d+ AMR ($n = 24$) and C4d– AMR ($n = 20$) vs. controls ($n = 20$) AMR– | Significantly more cAMR and graft loss on follow-up (between-group analysis [Mann–Whitney], not outcome analysis) | Very low |
| Orandi et al. (31) | Graft loss | Banff 2013 Methods do not clarify whether cases with cg are included | Retrospective | N = 217 patients with AMR; (142 clinical AMR, 77 subclinical) + controls (426 clinical, 231 subclinical); high proportion (63%) of HLA-incompatible transplants, so may not apply to conventional transplantation | Graft loss in subclinical AMR 2.15-fold (95% CI 1.19–3.91; $p = 0.012$) higher than for matched controls without AMR; graft loss clinical AMR 5.79-fold (95% CI 3.62–9.24; $p < 0.001$) higher than for matched controls without AMR | Low +1 (RR > 5 for clinical AMR) |
| Orandi et al. (32) | TG and graft loss | Banff 2013 aAMR and/or caAMR likely both included | Single-center, retrospective; all biopsies in 1st year post transplant (indication + protocol) | 51 C4d– and 156 C4d+ cases of AMR | TG risk same in C4d– and C4d+ but not vs. controls; 1-year and 2-years post-AMR graft survival: C4d– vs. controls: 2.56-fold (95% CI 1.08–6.05; $p = 0.033$); C4d+ vs. controls 3.70-fold (95% CI 2.47–5.54; $p < 0.001$); no difference between C4d– and C4d+ | Low |
| Everly et al. (33) | Graft loss | Altered Banff definition used: Banff 2003 AMR including suspicious (i.e., if ≥ 2 of the following present: DSA, histopathologic changes consistent with AMR and C4d+ staining in PTC ± other structures); do not specify active or chronic active | Retrospective | Patients with acute cellular rejection ($n = 30$) or AMR ($n = 30$) | Significantly worse survival in AMR ($p < 0.001$) | Low |
| Loupy et al. (34) | TG; 1/sCr and eGFR | Altered Banff definition used (Banff 1997 + addition of C4d– AMR); contains some cg/chronic cases at baseline | Retrospective | Pre-sensitized patients ($n = 54$) | Subclinical AMR at 3 months associated with interstitial fibrosis and tubular atrophy, TG and worse function at 1 year | Low |

*a, acute/active; AMR, antibody-mediated rejection; c, chronic; ca, chronic active; CI, confidence interval; DSA, donor-specific antibody; eGFR, estimated glomerular filtration rate; HLA, human leukocyte antigen; PTC, peritubular capillaries; sCr, serum creatinine; TG, transplant glomerulopathy.*

than the current threshold of glomerular basement membrane "double contours (incomplete or circumferential) in at least three glomerular capillaries by EM, with associated endothelial swelling and/or subendothelial electron-lucent widening" (19, 23)—the publication reported ~50% graft loss within 24 months after the index biopsy. There appears to be no difference in outcome between cases with C4d positivity and DSA negativity (qualifying as caAMR according to Banff 2017/2019) and TG cases with C4d negativity and DSA positivity (37) (qualifying as cAMR or caAMR if moderate MVI is present, according to Banff 2017/2019) (19, 20, 23).

To investigate associations between TG and other parameters, as well as with outcomes, using archetypal analysis, a retrospective study of 385 patients with TG identified five distinct immunological, histological, and functional profiles of TG that were associated with allograft failure (54). Another retrospective analysis of TG in 954 kidney transplant recipients (3744 biopsies) found that TG occurred in >75% of the patients in the absence of HLA-DSA, independent of HLA molecular mismatches; it represented a different phenotype that had lower levels of concomitant inflammation and graft loss compared with HLA-DSA+ TG (56). An additional recent retrospective study found that

**TABLE 3 |** Studies investigating associations between Banff diagnostic categories cAMR/caAMR and outcome; and/or investigating TG (37–54).

| References | Endpoint | Parameter investigated | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Lesage et al. (37) | Graft loss | GBM splitting | ~50% graft loss within 24 months after index biopsy; HR even after adjustment for sCr and proteinuria >5 vs. controls | Moderate |
| Wavamunno et al. (38) | Death-censored transplant survival | Ultrastructural | No difference in death-censored transplant survival between 7 patients with Banff cg ≥ 1 (Banff 1997) within the first 5 years post transplantation vs. 8 controls | Very low |
| Perkowska-Ptasinska et al. (39) | Transplant survival in subgroups of TG | Banff cg ≥ 1 Banff 2007 = Banff 1997 | 38/158 patients with TG lost transplant within 98 months (range, 3–215 months); no control cohort; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Shimizu et al. (40) | Graft loss in patients with a diagnosis of TG | Banff cg ≥ 1 Banff 2009 = Banff 1997 | 22% graft loss within observation time (time unclear); includes ABO-incompatible transplants; Unclear whether Banff chronic AMR fulfilled; no comparison with controls regarding outcome | Not applicable |
| Hayde et al. (41) | Death-censored transplant survival | cAMR compared to IFTA and TG (authors' own definitions; invalid criteria for TG: "by electron microscopy . . . electron-lucent widening of the subendothelial zone of the GBM, subendothelial accumulation of flocculent material, with or without a new subendothelial basement membrane layer") | cAMR associated with significantly lower graft survival compared with IFTA ($p = 0.01$) but not compared with TG | Low |
| Pefaur et al. (42) | Transplant survival | Unclear criteria for TG | Retrospective study, 3 patients; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Shimizu et al. (43) | Transplant survival | Banff cg ≥ 1 Banff 2007 = Banff 1997 | Retrospective study, 13 patients, no control group; 2/13 grafts lost; unclear observation time; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| John et al. (44) | Death-censored transplant survival | Banff cg ≥ 1 Banff 1997 | Retrospective study, 36 patients with TG, 5-years death-censored graft loss 16.7%; no control group; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Nair et al. (45) | Graft loss | Unclear criteria for TG that do not necessarily involve GBM splitting | Three patients with TG within first 6 months post transplantation; no control group; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Lopez Jimenez et al. (46) | Graft loss | Banff cg ≥ 1 Banff 1997 | Retrospective study, 30 patients with TG; 50% graft loss mean 25 ± 20 months post biopsy; no control group; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Kamal et al. (47) | Graft loss | Banff cg ≥ 1 Banff 2007 + 2009 = Banff 1997 | Retrospective study, 52 patients with TG, 17 (32%) with graft loss, median time to graft loss 16 months, no control cohort; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Dobi et al. (48) | Transplant survival | Banff cg ≥ 1a Banff 2013 | Retrospective analysis, 57 patients with TG; no control cohort; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Halloran et al. (49) | Death-censored transplant survival | Banff cg ≥ 1a Banff 2013 | Retrospective analysis, 27 patients with TG; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Toki et al. (50) | No outcome data (eGFR at time of biopsy) | Banff cg ≥ 1b Banff 2013 | Retrospective analysis, 127 patients with TG; extraction of outcome data on c/caAMR/TG not possible; no outcomes data | Not applicable |
| Courant et al. (51) | Death-censored transplant survival | cAMR according to Banff 2013 with DSA+ | Retrospective data, 9 patients with cAMR; extraction of outcome data on c/caAMR/TG not possible | Not applicable |

(Continued on following page)

**TABLE 3 |** (*Continued*) Studies investigating associations between Banff diagnostic categories cAMR/caAMR and outcome; and/or investigating TG (37–54).

| References | Endpoint | Parameter investigated | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Lubetzky et al. (52) | Transplant survival | Banff cg ≥ 1 Banff 1997; patients with TG and DSA+/MVI– | Retrospective analysis, 24 patients 50% graft loss in 3 years; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Sablik et al. (53) | Transplant survival | Banff cg ≥ 1b Banff 2015 | Retrospective analysis, 41 patients with caAMR; no control cohort; extraction of outcome data on c/caAMR/TG not possible | Not applicable |
| Aubert et al. (54) | Transplant survival | Banff 2009, 2011, 2013. Unclear if cg1a included | Retrospective analysis, 385 patients with TG; different immunological, functional and histological TG subtypes described; no comparison to control group without TG; extraction of outcome data on c/caAMR/TG not possible | Not applicable |

*a, acute/active; AMR, antibody-mediated rejection; c, chronic; ca, chronic active; DSA, donor-specific antibody; eGFR, estimated glomerular filtration rate; GBM, glomerular basement membrane; HR, hazard ratio; IFTA, interstitial fibrosis and tubular atrophy; MVI, microvascular inflammation; sCr, serum creatinine; TG, transplant glomerulopathy.*

proteinuria, C4d presence, and mesangial matrix expansion were important for outcome, while other histological markers (e.g., Banff Lesion Score cg) were not (57).

Because of repeated revisions to Banff criteria (including gene transcripts and the requirement for EM, to detect PTCML and early TG lesions), the incidence of caAMR is under-reported. No studies fulfill all criteria for this diagnosis according to Banff 2017 or have a sufficient follow-up to use strictly defined Banff caAMR as an endpoint (58).

We therefore recommend that clinical trials in kidney transplantation using caAMR as an endpoint or an inclusion criterion strictly adhere to Banff consensus criteria and report granular histological features of Banff Lesion Scores, to allow between-trial comparisons. Additional research is needed, in high- and conventional-risk transplantation scenarios that consider the effect of treating aAMR earlier, equally defined according to the strict Banff Classification.

## Suspicious for AMR Subcategories

A noteworthy change to the Banff Classification in 2017 was its omission of 'suspicious for aAMR' and 'suspicious for caAMR' categories (19). The most frequent reason leading to a diagnosis of 'suspicious for "aAMR" instead of "aAMR" was absence of evidence for DSA or C4d positivity (9).

Until 2019, no publication presented outcomes for patients with "suspicious for aAMR." Then, the evidence appeared, with a caveat, because the 123 DSA– patients with AMR included six patients with TG (Banff Lesion Score cg ≥ 1); irrespective of C4d status, outcomes for patients with histological features of AMR but without DSA were no different than for controls without AMR (59). Although there was a significant association between C4d status and DSA in this study, C4d and DSA were not interchangeable (accuracy of C4d deposition for DSA positivity was 59–65%) (59).

The literature offers even less information about the diagnostic subcategory of "suspicious for caAMR," eliminated from Banff in 2017. One study involving 21 DSA– patients showed an average transplant survival after diagnosis of 3.7 years (53).

Some DSA– cases "suspicious for AMR" could be explained by non-HLA antibodies. Without any hard evidence, standardized tests, or validated assay and cut-off value to screen for non-HLA-DSA, we do not recommend that non-HLA-DSA be considered in the diagnosis of AMR. Further research is needed before non-HLA antibodies can be included in the definition of endpoints for registration studies.

Overall, we do not recommend using cases in the "suspicious for" categories as endpoints.

# INDIVIDUAL HISTOPATHOLOGICAL FEATURES OF AMR AS ENDPOINTS

## "ATI in the Absence of Any Other Apparent Cause" as a Feature of AMR, in Conjunction with C4d Positivity and DSA

This section reports on two category 2 diagnoses that are separated in the Banff 2017/2019 Classification (19,21). Firstly, aAMR, where evidence of tissue injury is only "ATI in the absence of any other apparent cause" (ATI-AMR); to diagnose aAMR in such cases, C4d must be positive. Secondly, "C4d-staining without evidence of rejection": this is a subcategory of "antibody-mediated changes." Evidence relating to both entities is reviewed together, because the difference between them relates to presence or absence within the biopsy of the Banff additional diagnostic parameter "ATI in the absence of any other apparent cause." ATI has not been redefined since the 1995 Banff meeting; most transplant biopsies show a mild degree of ATI that might not qualify for ATI-AMR; at the lower end of the spectrum of ATI severity, the difference between ATI-AMR and C4d+ without evidence of rejection is tenuous. We are not aware of an evidence-based definition separating ATI-AMR and ATI of other causes.

Technically, according to the Banff classification, a biopsy that is C4d+ with DSA but with a reasonable other cause of ATI (e.g., ischemia/reperfusion injury) is not AMR, yet publications have not assessed for (or reported on) other causes of ATI. Some early reports on C4d staining date from before widespread recognition of the full spectrum of histological features of aAMR, therefore descriptions of poor outcomes for C4d+ cases "without features of rejection" must be handled cautiously.

The 2001 Banff meeting recognized a form of AMR with no or little inflammation, included in the list of category 2 diagnoses as "acute tubular necrosis-like minimal inflammation, C4d+." It was stated that "acute humoral rejection may be manifested only by ATI without other evidence of rejection (seen in 10% of cases)." The evidence cited (60) describes two cases of AMR where ATI-like changes were the sole feature. It is likely that inclusion of the ATI-AMR in Banff 2001 was based on the combined experience of meeting attendees, from an era when less sensitive pre-transplant evaluation for HLA antibodies created a population of accelerated/acute AMR with these features. Current data on the incidence of this histological variant, in both low- and high-risk transplantations, are lacking.

The 2007 Banff meeting (18) described a different subcategory of antibody-mediated changes that is now called "C4d staining without evidence of rejection" (19). It includes cases with C4d+ staining, but no features of activity or chronicity related to AMR (**Table 1**), and no features of TCMR or borderline changes. Banff 2017 further specified that there should be no evidence of increased expression of thoroughly validated gene transcripts or classifiers in biopsy tissue samples strongly associated with AMR (19). This diagnosis excludes cases with "ATI in the absence of any other apparent cause," although—as stated above—it is likely that mild ATI features are frequently observed. This category includes biopsies from recipients of ABO-incompatible transplants, in which it is associated with good outcomes (19), but also includes cases from recipients of ABO-compatible transplants, in which case its significance is unclear.

Additional publications investigating the link between ATI-AMR and C4d+ without evidence of rejection, with outcome data, are presented in **Table 4** (61–66). These studies provide low-quality evidence, but further research might have an impact on confidence in the estimate and could change the assessment. The data suggest that, in sensitized patients, C4d+ ATI (likely severe) in the early post-transplant phase could represent early AMR and be associated with graft loss (61–63), whereas the significance of C4d+ with mild ATI in later post-transplant biopsies is less clear. Some evidence suggests it is not strongly associated with future AMR or graft loss (64–66).

It is impossible to give a guideline recommendation because of inconsistent findings in the argument that C4d+ ATI without evidence of rejection is associated with increased risk of graft loss. In addition, there is no recent consensus definition of ATI, or of degrees of severity of ATI, or of what reasonably constitutes exclusion of other causes of ATI. Therefore, we recommend that the C4d+ ATI-only form of AMR and C4d+ without evidence of rejection subcategory of AMR should not be used as an efficacy measure in clinical trials. We also recommend that future research incorporates definitions of ATI and assessments of its severity, based on definitions agreed in the context of international collaborations (e.g., Banff Working Group for Rules and Dissemination). Such research should include both patients with preformed antibodies (sensitized) and non-sensitized patients, with representation of early and late post-transplant periods.

## Endarteritis

Endarteritis is also a feature of aTCMR that initially was not included in AMR definitions; this makes findings from early studies difficult to interpret for the given purpose (Banff Lesion

**TABLE 4 |** Studies investigating ATI-AMR and C4d+ without evidence of rejection (61–66).

| References | Endpoint | Definition of Banff phenotype | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Haas et al. (61) | AMR | C4d+ in early post-reperfusion biopsies | Predicts future AMR ($n$ = 2 positive crossmatch patients with later AMR) | Low |
| Djamali et al. (62) | AMR | C4d+ in early post-reperfusion biopsies; mild to moderately sensitized transplant recipients | Predicts future AMR | Low |
| Kikic et al. (63) | Graft loss | Biopsies with C4d; 42% of patients in the C4d+ group were pre-sensitized; mean time to biopsy in C4d+ group 0.75 mo | C4d associated with graft loss independently of presence of histological features of AMR; HR 1.85 ($p$ < 0.0001) | Low |
| Nickeleit et al. (64) | Benefit from antirejection therapy | C4d+ with mild allograft dysfunction and no histological evidence of rejection | C4d+ with mild allograft dysfunction and no histological evidence of rejection does not benefit from antirejection therapy | Low |
| Dickenmann et al. (65) | Improved function after treatment | C4d+ biopsies without other histopathological features of AMR | Function improves in this group after treatment | Low |
| Dominy et al. (66) | AMR | C4d+ without evidence of rejection; mild ATI at most | Rather than histological features or DSA, transcript analysis for AMR signature distinguishes minority at risk of subsequent AMR | Low |

*AMR, antibody-mediated rejection; ATI, acute tubular injury; DSA, donor-specific antibody; HR, hazard ratio.*

Score v in "acute TCMR," "mixed acute TCMR-AMR," "pure AMR"). Although endarteritis was described as a risk factor for graft loss (67), there are insufficient published data on endarteritis in pure aAMR as an isolated histopathological finding to recommend its use as an AMR-related endpoint.

## Acute TMA in the Absence of Any Other Cause

Banff acknowledges that TMA can have a variety of causes in kidney transplant recipients (e.g., recurrent disease, infection, antiphospholipid antibodies, medication toxicity). During the 2015 Banff meeting, a working group was formed (68) to help with histopathological characterization of TMA in kidney transplantation. This group aimed to guide the development of precise diagnostic algorithms, including the creation of rules on how other apparent causes could be excluded, allowing for a *bona fide* diagnosis of AMR-associated TMA. In some patients with *dn*TMA, an underlying genetic defect in complement regulation might be relevant, although only one case series suggested this (69).

We are unaware of sufficient published data about the outcomes of adequately investigated cases of AMR-associated TMA. The largest case series describes 33 patients with TMA and C4d positivity, 40% of whom experienced transplant failure within 2 years of diagnosis (70). Since C4d positivity in peritubular capillaries and medullary vasa recta is extremely rare in native kidneys with TMA (71), this combination of findings can be considered "AMR-associated TMA," as is currently the case according to Banff 2017 (19). However, the problem persists of excluding other causes of TMA. Nevertheless, for reasons outlined above, we would not encourage the use of TMA as isolated histopathological finding as an efficacy measure for clinical trials, in a context that does not meet the Banff diagnostic criteria for a full AMR phenotype. Nor is there enough evidence to recognize "acute TMA in the absence of any other cause" as a sufficiently robust criterion for aAMR.

## Microvascular Inflammation

MVI is the main histological feature indicating activity in aAMR and caAMR. The Banff criteria for AMR use cut-off values of MVI >0 and >1, respectively, to establish C4d+ and C4d− aAMR; these values were established by consensus, based on published evidence (19). MVI above a certain threshold in diagnostic biopsies is an independent predictor of graft loss and chronic lesions (**Table 5**) (72–77), although the quality of evidence is low. Moreover, low interobserver agreement in the exact grading of the underlying g and ptc lesions (24) suggests caution when using this parameter as an endpoint in studies.

Based on the low-quality evidence that MVI is an independent predictor of graft loss, we cautiously recommend that the MVI score is used as an efficacy marker for clinical trials in kidney transplantation. We also recommend that further research is undertaken to confirm the effect of MVI on outcome, in prospective randomized controlled trials, with granular histological data for Banff Lesion Scores and DSA.

## C4d Positivity

There are caveats to the prognostic value of C4d status: thresholds for positivity scoring differ, depending on antibody and study. For example, the monoclonal antibody used on frozen tissue is particularly sensitive; therefore >10% of PTC must be positive, whereas for the polyclonal anti-C4d antibody used on formalin-fixed paraffin-embedded tissue samples, any percentage and

**TABLE 5 |** Studies investigating the association between MVI and outcomes (72–77).

| References | Endpoint | Predictor | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Haas and Mirocha (72) | TG | MVI + endothelial lesions on EM | Indication biopsies (DSA at time of biopsy): MVI + endothelial lesions on EM associated with TG | Low |
| Bagnasco et al. (73) | TG | g | Patients with pre-transplant DSA (deceased donors including ABO-incompatible donors): g in protocol + indication biopsies associated with TG ($p < 0.0001$) | Low |
| Einecke et al. (74) | Graft loss | MVI | In multivariate analysis (indication + protocol biopsies, DSA post-transplant, living + deceased donors), graft failure correlated with MVI and scarring, but C4d staining was not significant | Low |
| Sis et al. (75) | Graft loss | MVI | Indication biopsies (anti-HLA antibodies at time of biopsy, living + deceased donors): g + ptc predicted graft failure independently of time, C4d and transplant glomerulopathy ($p < 0.001$) | Low |
| Verghese et al. (76) | Graft loss | MVI | Retrospective, no data on DSA, includes mixed TCMR-AMR indication biopsies. In indication biopsies carried out <1 year post-transplant, MVI associated with decreased death-censored graft survival, independent of the presence of C4d ($p = 0.005$) | Low |
| de Kort et al. (77) | Graft loss | MVI | Retrospective cohort study indication biopsies of patients with *dn*DSA: severe MVI >21-fold increased risk of graft failure (95% CI 2.5–180.0; $p = 0.005$), while C4d positivity on indication biopsy lost significance | Low |

*AMR, antibody-mediated rejection; CI, confidence interval;* dn, de novo*; DSA, donor-specific antibody; EM, electron microscopy; HLA, human leukocyte antigen; MVI, microvascular inflammation; TCMR, T cell-mediated rejection; TG, transplant glomerulopathy.*

intensity of PTC positivity sufficiently describes a biopsy as C4d+ (23).

Literature findings related to the potential prognostic value of C4d status are varied, likely because of the dynamic process (**Table 6**) (3, 63, 78, 79). A large body of evidence indicates that MVI is a better prognostic factor than C4d (74–77). C4d positivity as an isolated histopathological finding therefore cannot be recommended as an efficacy measure for clinical trials in kidney transplantation.

## Transplant Glomerulopathy

As TG is the main feature indicating chronicity in the diagnosis of caAMR and cAMR, much of the evidence indicating that this feature is an indicator of outcome has been covered above (**Table 3**). As with all individual histological lesions, moderate interobserver agreement in the graded scoring (24) suggests trial results should be interpreted cautiously.

In 55 patients with TG (Banff Lesion Score cg ≥ 1b) (19) there was a high risk of death-censored transplant survival in a multivariate analysis (HR 6.2; 95% CI 2.5–14.7; $p < 0.0001$) (80). Similar results were obtained in another multivariate analysis of 77 indication biopsies (HR 2.40; 95% CI 1.25–4.60; $p < 0.01$); this study used a lower threshold (Banff Lesion Score cg > 0) (21), equivalent to Banff 2017 (19, 23), but did not mention how many biopsies were examined with EM (51).

Applying Racusen's criterion for defining glomerular basement membrane splitting, mentioned above (2), which has a higher threshold for TG than current criteria (19,23), Torres *et al.* identified ~50% graft loss within 3 years after the index biopsy (81). Using the same threshold for TG, a retrospective study found graft loss in 2/12 patients with isolated TG in the absence of sufficient MVI, C4d positivity, or DSA positivity at the time of biopsy; notably, their definition did not necessarily exclude caAMR according to Banff 2017 (23, 68). While glomerular basement membrane splitting is a prerequisite for diagnosing TG as a manifestation of cAMR, it is by no means specific to AMR and can arise in different conditions—some of which are recognized by Banff—including TMA of causes other than AMR, *dn* or recurrent glomerulonephritis (23), hepatitis C virus infection (82), or hypertensive glomerulopathy (83). We recommend that further research is performed to establish the causes and impact of isolated TG that does not fulfill criteria for cAMR or caAMR. TG as an isolated histopathological finding cannot be recommended as an efficacy measure for clinical trials in kidney transplantation.

## Peritubular Capillary Basement Membrane Multilayering

Normal peritubular capillaries have a single basement membrane under the endothelial cell, and PTCML is characterized by an increase in basement membrane layers. Low levels of PTCML are seen in several conditions, whereas severe PTCML is a defining feature of chronicity in the Banff definition of AMR. Severe PTCML is characterized as seven or more layers of basement membrane in at least a single cortical peritubular capillary, and five or more layers in at least two additional capillaries.

Currently, PTCML is only diagnosed in transplantation biopsies by EM evaluation, which limits its use to centers sufficiently resourced to undertake such examinations. Even within EM-capable centers, this diagnostic method may be reserved for cases for which there is an indication [defined in Banff 2013 (21)]. There is therefore an inherent bias in reports investigating PTCML, which generally do not involve systematic assessment of all biopsies.

The limited number of observational studies investigating the link between PTCML and outcome (**Table 7**) (48, 84–86) provide low-quality evidence: further research is likely to have an impact on confidence in the estimate and may indeed change it. Although there is consistent evidence that PTCML is associated with future TG and increased risk of graft loss, it is impossible to give a guideline recommendation or consensus-based statement, because studies use different methodologies. Therefore, we recommend that PTCML as an isolated histopathological finding is not used as efficacy measure for clinical trials. We also recommend that future research

**TABLE 6 |** Transplantation studies that feature C4d and outcomes (3, 63, 78, 79).

| References | Endpoint | Predictor | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Naesens et al. (3) | Graft loss | Composite | Retrospective (indication biopsies, no info on DSA): C4d, TG, ongoing interstitial inflammation, dn/ recurrent glomerular disease, IFTA significantly and independently associated with post-biopsy graft survival (MVI highly significant in univariate, but not retained in final multivariate model) | Low |
| Kikic et al. (63) | Graft loss | C4d | C4d associated with graft loss independently of the presence of histological features of AMR | Low |
| Sapir-Pichhadze et al. (78) | Graft loss | C4d | Systematic review (3492 abstracts: 3485 indication and 868 protocol biopsies). C4d+ associated with inferior allograft survival compared with DSA or histopathology alone | Low |
| Matas et al. (79) | Graft loss | C4d | Cross-sectional (retrospective) cohort (indication biopsies, DSA at time of biopsy, living + deceased donors): C4d−/DSA- recipients had significantly better (and C4d+/DSA+ worse) death-censored graft survival than other groups. C4d+/DSA− and C4d−/DSA+ had similar intermediate death-censored graft survival | Low |

*AMR, antibody-mediated rejection;* dn, de novo*; DSA, donor-specific antibody; IFTA, interstitial fibrosis and tubular atrophy; MVI, microvascular inflammation; TG, transplant glomerulopathy.*

**TABLE 7 |** Studies that feature PTCML assessment and outcomes (48, 84–86).

| References | Endpoint | Definition of PTCML | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Einecke et al. (84) | Graft loss | One PTC with ≥5 basement membrane layers | In non-selected transplant population, 1 PTC with ≥5 basement membrane layers predictive of graft loss in multivariate analysis (HR 1.98, $p = 0.01$) | Low |
| Roufosse et al. (85) | TG | Numbers of PTC with ≥3 and ≥5 basement membrane layers | Risk of TG increases with increasing numbers of PTC with ≥3 and ≥5 basement membrane layers | Low + 1 ('dose–response' gradient) |
| de Kort et al. (86) | Graft loss | Three PTC with ≥5 basement membrane layers | In patients with *dn*DSA, 3 PTC with ≥5 basement membrane layers associated with increased graft loss ($p = 0.016$) | Low |
| de Kort et al. (86) | TG | Mean basement membrane layer count >2.5 | Mean PTCML count >2.5 associated with increased risk of TG ($p = 0.001$); progressors to >2.5 associated with more TG | Low |
| Dobi et al. (48) | Graft loss | PTC circ score ≥3 | In patients with cAMR, PTC circ ≥3 predicts graft loss | Low |

*AMR, antibody-mediated rejection; c, chronic;* dn, de novo*; DSA, donor-specific antibody; HLA, human leukocyte antigen; HR, hazard ratio; PTC, peritubular capillary; PTCML, peritubular capillary basement membrane multilayering; TG, transplant glomerulopathy.*

incorporates methods of counting basement membrane layers that are agreed in the context of international collaborations (e.g., Banff Working Group for Electron Microscopy), assess non-selected populations of transplant biopsies, and utilize clinically meaningful scoring systems that predict graft loss and cAMR development.

## Transplant Vasculopathy

The definition of transplant vasculopathy as evidence of AMR chronicity remains ambiguous in the Banff Classification (19, 23). Consequently, using the sole finding of transplant vasculopathy cannot be encouraged as an efficacy measure for clinical trials. We are unaware of any studies reporting outcomes for patients with this criterion for AMR chronicity. Older publications discussing the impact of Banff Lesion Score cv without specification of morphological details of this finding (88) are unhelpful, because this score can be influenced by factors other than AMR and can be ≥ 1 even in implantation biopsies (donor-derived).

## Increased Transcripts or Transcript Sets Strongly Associated with AMR

Increased expression of thoroughly validated gene transcripts/classifiers in biopsy tissue strongly associated with AMR provides evidence of current/recent antibody–tissue interactions, according to the Banff 2017/2019 definition of AMR (19, 22). Notably, many publications relating to transcript analysis do not distinguish between TCMR and AMR, which limits the studies that can be included here. Retrospective investigations of the link between gene transcripts/classifiers strongly associated with AMR and outcome also provide low-quality evidence (**Table 8**) (4, 47, 88–91). However, the INTERCOM study prospectively analyzed 300 transplantation biopsies (264 patients) and found that assigning an AMR score based on molecular analysis identified signs of AMR in 41% of biopsies where AMR had not been suspected: the score also showed a better correlation with graft failure than conventional assessments (92). The MMDx Kidney study group also prospectively collected microarray data from >1200 transplant biopsy samples and found that precision microassessment enabled six archetypes to be generated (from no rejection through TCMR and all stages of AMR) (93). Further research could have an impact on confidence in the estimate and might change it.

Although there is consistent evidence that gene transcripts/classifiers strongly associated with AMR are associated with graft loss (and in some cases, the evidence comes from multivariate analyses with validation groups), it is impossible to give a guideline recommendation or consensus-based statement. Different gene sets/classifiers are used across the studies, with no unifying set of genes agreed on for future validation in prospective research. Also, to our knowledge, no transplant centers have clinical validation for use of transcript analysis for AMR, especially for improving the prediction of graft outcome. Consequently, we recommend that gene transcripts/classifiers strongly associated with AMR are not used as efficacy measures for clinical trials. Future research in the context of international collaborations on agreed gene sets/classifiers (e.g., Banff Working Group for Molecular Pathology) should assess non-selective populations of transplant biopsies and determine clinically meaningful molecular scoring systems that predict cAMR development and graft failure. These studies should include multivariate analyses in combination with traditional clinical, histopathological, or immunogenetic parameters.

## Subclinical AMR Including Incomplete Phenotypes

**Table 9** lists publications that describe subclinical AMR in protocol biopsies, including incomplete phenotypes, and all studies linking Banff diagnostic categories and subcategories to

**TABLE 8 |** Studies that feature 'Evidence of gene transcripts/classifiers strongly associated with AMR' and outcomes (4,47,88–91).

| References | Endpoint | Definition of molecular marker | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Sellares et al. (4) | Graft loss | AMR classifier | Retrospective cohort, 315 patients; AMR classifier predicts graft loss in Cox multivariate analysis | Low |
| (87) Loupy et al. (88) | Graft loss and progression of chronic injury | AMR molecular score and endothelial DSA-selective transcripts | 2 cohorts (principal $n = 74$, validation $n = 54$) with cases of AMR in 1st year after transplant (early AMR); AMR Molecular Score (HR 2.22; 95% CI 1.37–3.58; $p = 0.001$) and endothelial donor-specific antibody-selective transcripts (HR 3.02; 95% CI 1.00–9.16; $p < 0.05$) independently associated with increased risk of graft loss | Low |
| Yazdani et al. (89) | Graft loss | Differential expression of 503 unique genes in AMR, with significant enrichment of NK cell pathways | Retrospective cohort, with validation in external cohort for outcome analysis; microarray transcriptomic data from case–control study ($n = 95$) to identify genes differentially expressed in AMR; multivariate Cox analysis: NK cell gene signature predicted graft loss better than ($p < 0.001$), and independent of, the diagnosis of rejection according to Banff ($p = 0.039$) | Low |
| Sis et al. (90) | Graft loss | ENDATs | Retrospective cohort of indication biopsies validated in independent set; microarray analysis for ENDATs. Many individual ENDATs were increased in AMR and predicted graft loss; high ENDAT score in patients with DSA predicts graft loss (but no increase in graft loss if DSA– and ENDAT+) | Low |
| Dominy et al. (91) | Graft loss | *Sh2D1b* and *Mybl* score | Retrospective cohort of 57 biopsies from patients with AMR or normal surveillance biopsies; 2-gene signature predicts graft loss in whole group and within DSA+ group | Low |
| Kamal et al. (47) | Graft loss; no formal outcome analysis | Various gene expression levels | Retrospective cohort of patients with TG; significantly increased levels endothelial cell–associated transcripts, gene transcripts associated with complement cascade, interleukins and their receptors, and granulysin in patients with graft loss | Very low |

*AMR, antibody-mediated rejection; CI, confidence interval; DSA, donor-specific antibody; ENDAT, endothelial cell–associated transcript; HR, hazard ratio; NK, natural killer; TG, transplant glomerulopathy.*

outcomes in protocol biopsies (34, 80, 94–99). Subclinical AMR diagnosed in protocol biopsies is associated with subsequent chronic kidney injury, impaired graft function, and impaired graft survival, but whether treatment of subclinical AMR diagnosed in protocol biopsies improves graft outcomes is not proven. The quality of evidence is not high.

A literature search for studies evaluating the frequency of subclinical AMR management showed that ~60% of patients received treatment, usually with antibody-targeted therapies. Again, national variations were observed. In Paris, 57% of patients with subclinical AMR received antirejection therapy (36) while US centers treated subclinical AMR more aggressively than elsewhere (100); centers in Canada (101) and Belgium (59) treated this presentation very selectively. Differences may also relate to whether centers perform high-risk transplantations and the timing of the post-transplant biopsy. Early (e.g., 1- or 3-month) post-transplant subclinical AMR in patients at high immunological risk may have different outcomes than late (e.g., ≥1-year) post-transplant subclinical AMR in patients with *dn*DSA. Given that subclinical AMR in protocol biopsies appears to be associated with impaired graft survival, but protocol biopsies are not universally performed, and the management of subclinical AMR is heterogeneous, it is unsurprising that consensus documents do not provide guidance (28, 102). We consider that identifying AMR in protocol biopsies could be a clinically meaningful endpoint as an independent predictor of graft loss; but in the absence of high-quality evidence and uncertainty about the effect of treatment, we remain cautious. The priority should be to agree good definitions for the phenotypes and endpoints of AMR that are clinically meaningful in kidney transplantation studies. We also recommend that further research investigates the role of subclinical AMR in graft failure.

## Restricted Definition of Banff Classification of AMR for Use as Endpoint

Based on the evidence presented above, we propose a restricted definition of the Banff phenotypes of AMR, if used as an endpoint in interventional trials (**Table 10**) (19, 22).

**TABLE 9 |** Studies investigating outcomes in cases with subclinical AMR, including incomplete phenotypes (34, 80, 94–99).

| References | Endpoint | Definition of predictor | Findings | Level of evidence (grade) |
|---|---|---|---|---|
| Loupy et al. (34) | GFR, TG, IFTA | Subclinical AMR | Patients with subclinical AMR at 3 mo had at 1 year: Higher rate of IFTA (100% vs. 33.3%; $p < 0.01$) Higher rate of TG (43% vs. 0%; $p = 0.02$) Lower mGFR (39.2 ± 13.9 vs. 61.9 ± 19.2 ml/min/1.73 $m^2$; $p < 0.01$) | Low |
| Lerut et al. (94) | PTCML, cAMR | PTC | Protocol biopsies with ptc at 3 mo associated with PTCML ($p < 0.0001$)/cAMR ($p = 0.0002$) at 1 year | Low |
| Haas et al. (95) | CAN score (cg + ci + ct + cv) | Subclinical AMR | Subclinical AMR (stable SCr, PTC, diffuse PTC C4d, positive DSA) during 1st year post transplantation associated with higher increase in CAN score in follow-up biopsies 335 ± 248 (SD) days later (3.5 ± 2.5 vs. 1.0 ± 2.0; $p = 0.01$) | Low |
| Loupy et al. (96) | cAMR | MVI + class II DSA | Multivariate analysis demonstrated that presence of MVI and anti-HLA class II DSA at 3 mo was associated with a 4-fold increased risk of progression to cAMR independently of C4d ($p < 0.05$) | Low |
| Cosio et al. (97) | Graft loss | 'cAMR' (= cg > 0 or MVI≥2) | 4.2% of protocol biopsies at 1 year showed cAMR; risk of death-censored graft survival HR 12.6 (95% CI 6.58–24.3; $p < 0.0001$) | Moderate |
| Gloor et al. (80) | GFR, proteinuria | TG | Prognosis of subclinical TG was equally poor as TG diagnosed with graft dysfunction, with progressive worsening of histopathologic changes and function | Low |
| Papadimitriou et al. (98) | | NR | Indication + protocol biopsies (concurrent DSA): More incomplete phenotype in protocol than in indication biopsies Persistence/worsening of AMR in a subsequent biopsy occurred in 38.2% of cases independently of strength of AMR findings in 1st biopsy (e.g., progression to cAMR occurred also in cases with suspicious or non-diagnostic findings) | Low |
| Tsuji et al. (99) | cAMR | MVI | MVI in protocol biopsies at 3 mo correlates with later development of cAMR ($p = 0.03$) | Low |

*AMR, antibody-mediated rejection; c, chronic; DSA, donor-specific antibody; GFR, glomerular filtration rate; HLA, human leukocyte antigen; IFTA, interstitial fibrosis and tubular atrophy; m, mean; MVI, microvascular inflammation; NR, not reported; PTC, peritubular capillary; PTCML, peritubular capillary basement membrane multilayering; sCr, serum creatinine; SD, standard deviation; TG, transplant glomerulopathy.*

## CONCLUSIONS

- Evidence relating to the relationship between AMR and outcomes is largely based on retrospective analyses that do not utilize the strict, most recent Banff categories of AMR, but instead investigate individual features of AMR, combinations of individual features of AMR, or combined Banff categories (such as combining aAMR and caAMR).
  - Strongest evidence for associations between individual features and impaired graft outcome is noted for MVI score ≥2 (if borderline changes, aTCMR or infection are present, g + ptc>1 is not sufficient and g > 1 is required) and cg>10% (>10% of the most severely affected glomerulus).
  - Together with presence of HLA-DSA, these parameters should be the basis for AMR endpoints, acknowledging their limitations (lack of specificity, between-study heterogeneity in definitions used, and high interobserver variability).
- Based on evidence for association between individual features of AMR and outcome, AMR diagnosed in indication or protocol biopsies should be considered as a primary endpoint in clinical trials for kidney transplantation.
- However, based on available evidence, we suggest refinement of the Banff 2017 definition for AMR diagnosis to the following three AMR-related endpoints:
  - Restricted aAMR, defined by the conservative threshold of at least moderate MVI (g + ptc≥2 with g ≥ 1 in the presence of aTCMR, caTCMR, or borderline changes) and DSA positivity (anti-HLA antibodies) with or without C4d positivity (C4d ≥ 1 on paraffin tissue or ≥2 on frozen tissue).
  - Restricted caAMR, defined by the conservative threshold of cg ≥ 1 according to Banff 2011 (≥10% of the glomerular capillary walls in the most severely affected glomerulus involved) plus at least moderate MVI (g + ptc≥2 with g ≥ 1 in the presence of aTCMR, caTCMR, or borderline changes) and DSA positivity (anti-HLA antibodies) with or without C4d positivity (C4d ≥ 1 on paraffin-embedded tissue; ≥2 on frozen tissue).
  - Restricted cAMR, defined by the conservative threshold of cg ≥ 1 according to Banff 2011 (≥10% of glomerular capillary walls in most severely affected glomerulus involved) and current or past DSA positivity (anti-HLA antibodies) with or without C4d positivity (C4d ≥ 1 on paraffin-embedded tissue; ≥2 on frozen tissue).
- Other features of AMR used in Banff AMR definitions (ATI in the absence of any other cause; TMA; Banff Lesion Score v ≥ 1; increased transcripts associated with AMR; cg<10%; PTCML;

**TABLE 10 |** Restricting the Banff classification for AMR for the purpose of endpoints in clinical trials, based on the evidence reviewed (19, 22).

| Banff 2017 Category 2: Antibody-mediated changes | Restricted definition of AMR for use as primary endpoint |
|---|---|
| **Active AMR; all criteria must be met for diagnosis** | **Active AMR; all criteria must be met for diagnosis** |
| 1. Histologic evidence of acute tissue injury, including ≥1 of the following | 1. Histologic evidence of acute tissue injury |
| • MVI (g > 0 and/or ptc>0), in the absence of recurrent or *de novo* glomerulonephritis, although in the presence of aTCMR, borderline infiltrate, or infection, ptc≥1 alone is not sufficient and g must be ≥ 1 | • At least moderate MVI (g + ptc≥2 with g ≥ 1 in the presence of aTCMR, caTCMR, or borderline changes, or infectious disease of the transplant) |
| • Intimal or transmural arteritis (v > 0) | • — |
| • Acute TMA the absence of any other cause | • — |
| • Acute tubular injury, in the absence of any other apparent cause | • — |
| 2. Evidence of current/recent antibody interaction with vascular endothelium, including one or more of the following | 2. Evidence of current/recent antibody interaction with vascular endothelium |
| • Linear C4d staining in peritubular capillaries (C4d2 or C4d3 by IF on frozen sections, or C4d > 0 by IHC on paraffin-embedded sections) | • At least MVI g + ptc≥2 with g ≥ 1 in the presence of aTCMR, caTCMR or borderline changes, or infectious disease of the transplant), identical to criterion 1 for aAMR |
| • At least moderate MVI ([g + ptc]≥2) in absence of recurrent/*dn* glomerulonephritis, although in the presence of aTCMR, borderline infiltrate, or infection, ptc≥2 alone is not sufficient and g must be ≥ 1 | • With or without C4d positivity (C4d ≥ 1 on paraffin tissue or ≥2 on frozen tissue) |
| • Increased expression of gene transcripts/classifiers in the biopsy tissue strongly associated with AMR, if thoroughly validated | • — |
| 3. Serologic evidence of donor-specific antibodies (DSA to HLA or other antigens). C4d staining or expression of validated transcripts/classifiers as noted above in criterion 2 may substitute for DSA; however thorough DSA testing, including testing for non-HLA antibodies if HLA antibody testing is negative, is strongly advised whenever criteria 1 and 2 are met | 3. DSA positivity (anti-HLA antibodies) |
| **Chronic active AMR; all criteria must be met for diagnosis** | **Chronic active AMR; all criteria must be met for diagnosis** |
| 1. Morphologic evidence of chronic tissue injury, one or more of the following | 1. Morphologic evidence of chronic tissue injury, including ≥1 of the following |
| • TGA (cg > 0) if no evidence of cTMA or chronic recurrent/*dn* glomerulonephritis; includes changes evident by EM alone (cg1a) | • cg ≥ 1 according to Banff 2011 (≥10% of the glomerular capillary walls in the most severely affected glomerulus involved) |
| • Severe peritubular capillary basement membrane multilayering (requires EM) | • — |
| • Arterial intimal fibrosis of new onset, excluding other causes; leukocytes within the sclerotic intima favor cAMR if there is no prior history of TCMR, but are not required | • — |
| 2. Identical to criterion 2 for aAMR, above | 2. Identical to criterion 2 for aAMR, above |
| 3. Identical to criterion 3 for aAMR, above, including strong recommendation for DSA testing whenever criteria 1 and 2 are met | 3. DSA positivity (anti-HLA antibodies) |
| **Chronic AMR; all criteria must be met for diagnosis** | **Chronic AMR; all criteria must be met for diagnosis** |
| 1. Morphologic evidence of chronic tissue injury, including ≥1 of the following | 1. Morphologic evidence of chronic tissue injury |
| • Transplant glomerulopathy (cg > 0) if no evidence of cTMA or chronic recurrent/*dn* glomerulonephritis; includes changes evident by EM alone (cg1a) | • Cg ≥ 1 according to Banff 2011 (≥10% of the glomerular capillary walls in the most severely affected glomerulus involved) if no evidence of TMA of any other cause or recurrent or *dn* glomerulopathy |
| • Severe PTCML (requires EM) | • — |
| 2. Absence of evidence of current/recent antibody interaction with the endothelium (criterion 2 for active AMR, above) | 2. Absence of evidence of current/recent antibody interaction with the endothelium (criterion 2 for aAMR, above) |
| 3. Prior documented diagnosis of a or caAMR or documented prior evidence of DSA | 3. Prior documented diagnosis of a or caAMR or documented prior evidence of DSA |

*a, acute/active; AMR, antibody-mediated rejection; ca, chronic active;* dn, de novo*; DSA, donor-specific antibody; EM, electron microscopy; HLA, human leukocyte antigen; IF, immunofluorescence; IHC, immunohistochemistry; MVI, microvascular inflammation; PTCML, peritubular capillary basement membrane multilayering; TCMR, T cell-mediated rejection; TMA, thrombotic microangiopathy.*

arterial intimal fibrosis of new onset; DSA− cases) show less-robust evidence than MVI score ≥2 and cg>10%.
  - In isolation, without the other features of AMR described above, these features should not be considered as efficacy endpoints for clinical trials.
- The use of histology as endpoint for studies after kidney transplantation needs to consider that histological scoring reproducibility is at best moderate.
- There is a clear need for additional investigations of outcomes for all features and all categories of AMR.
  - Any such studies should follow the Banff 2017 recommendations on best practice for pathology endpoints in clinical trials (19), in particular involving pathologists in clinical trial design, use of a panel of pathologists for grading with a defined adjudication mechanism, granular scoring and reporting of histological data as continuous parameters and, where possible, maintaining a digital archive of pathology slides to facilitate external validation; use of data lumped into arbitrarily defined 'AMR' is discouraged.

### Scientific Advice from the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA) Regarding These Conclusions

- The CHMP recognized the issues in defining AMR.
- The CHMP welcomed and endorsed the suggestion to initiate a discussion on the use of the Banff classification as a tool to define AMR as an endpoint in clinical trials, in addition to its diagnostic and research value.
- The rationale behind the restricted definitions of aAMR and caAMR for use as primary endpoints was well received by the CHMP.
  - For this to happen, evidence-based classification and state-of-the-art, transparent, and standardized review processes of scientific data are required to demonstrate the usefulness of the restricted Banff definitions for AMR.

## AUTHOR CONTRIBUTIONS

This article is one of several papers developed from the Broad Scientific Advice request, submitted to the EMA/CHMP by ESOT in 2020: interactions between the EMA and ESOT regarding this request began in 2016. For the present article, through discussions, working groups on histological and functional endpoints in kidney transplantation developed the ESOT position on the core question 'Does CHMP agree with the updated definitions of rejection and their potential use as primary endpoints in studies of kidney transplantation?'. The Center for Evidence in Transplantation provided support with data extraction requests: these literature searches formed the basis of evidence used in the advice request and the present article. Input into the working groups' outputs was provided from all ESOT members involved in the advice request process.

The present article was adapted by MN from the final Broad Scientific Advice request submission (June 2020), documentation from the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020). The first draft of the article was developed by JUB and CR and reviewed by DS and MR by e-mail. The article was finalized and approved by all co-authors before submission.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

CR is supported by the National Institute for Health Research (NIHR) Biomedical Research Centre based at Imperial College Healthcare NHS Trust and Imperial College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health. CR's research activity is made possible with generous support from Sidharth and Indira Burman. JUB consults for Sanofi. MR has received lecture fees from Astellas and Chiesi; and research grant support (paid to institution) from Astellas and Chiesi for investigator-initiated studies. GB has received honoraria and/or research funding from Astellas, CareDx, CSL Behring, Fresenius, Hansa, Neovii, and Vitaeris. KB has received honoraria and/or research funding from Alexion, Astellas, Bristol Myers Squibb, Chiesi, Fresenius, Hansa, Hexal, Merck, Novartis, Otsuka, Pfizer, Roche, Sandoz, Siemens, and Veloxis. DG reports consultancy agreements and scientific boards for AstraZeneca, BMS, Hansa, and Sanofi. LH reports speaker fees from Astellas, consultancy and research support from Chiesi, consultancy for Novartis, and research support from Sandoz. RO has received grants/research support from Amgen, Astellas, and Chiesi; and speakers' bureaux/honoraria from Amgen, Astellas, Chiesi, Hansa, Neovii, Novartis, and Teva. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. European Medicines Agency. Clinical Investigation of Immunosuppressants for Solid Organ Transplantation (2008). CHMP/EWP/263148/06. Available at: https://www.ema.europa.eu/en/clinical-investigation-immunosuppressants-solid-organ-transplantation (Accessed October 18, 2021).
2. Racusen LC, Solez K, Colvin RB, Bonsib SM, Castro MC, Cavallo T, et al. The Banff 97 Working Classification of Renal Allograft Pathology. *Kidney Int* (1999) 55:713–23. doi:10.1046/j.1523-1755.1999.00299.x

3. Naesens M, Kuypers DRJ, De Vusser K, Evenepoel P, Claes K, Bammens B, et al. The Histology of Kidney Transplant Failure. *Transplantation* (2014) 98: 427–35. doi:10.1097/tp.0000000000000183
4. Sellarés J, De Freitas DG, Mengel M, Reeve J, Einecke G, Sis B, et al. Understanding the Causes of Kidney Transplant Failure: the Dominant Role of Antibody-Mediated Rejection and Nonadherence. *Am J Transplant* (2012) 12:388–99. doi:10.1111/j.1600-6143.2011.03840.x
5. El-Zoghby ZM, Stegall MD, Lager DJ, Kremers WK, Amer H, Gloor JM, et al. Identifying Specific Causes of Kidney Allograft Loss. *Am J Transplant* (2009) 9:527–35. doi:10.1111/j.1600-6143.2008.02519.x
6. Chand S, Atkinson D, Collins C, Briggs D, Ball S, Sharif A, et al. The Spectrum of Renal Allograft Failure. *PLoS One* (2016) 11:e0162278. doi:10.1371/journal.pone.0162278
7. Mayrdorfer M, Liefeldt L, Wu K, Rudolph B, Zhang Q, Friedersdorff F, et al. Exploring the Complexity of Death-Censored Kidney Allograft Failure. *J Am Soc Nephrol* (2021) 32:1513–26. doi:10.1681/asn.2020081215
8. Chehade H, Pascual M. The challenge of Acute Antibody-Mediated Rejection in Kidney Transplantation. *Transplantation* (2016) 100:264–5. doi:10.1097/tp.0000000000000959
9. Wehmeier C, Amico P, Hirt-Minkowski P, Georgalis A, Höenger G, Menter T, et al. Acute Rejection Phenotypes in the Current Era of Immunosuppression: a Single-center Analysis. *Transplant Direct* (2017) 3: e136. doi:10.1097/txd.0000000000000650
10. Loupy A, Lefaucheur C. Antibody-mediated Rejection of Solid-Organ Allografts. *N Engl J Med* (2018) 379:1150–60. doi:10.1056/nejmra1802677
11. Karahan GE, Claas FHJ, Heidt S. Technical Challenges and Clinical Relevance of Single Antigen Bead C1q/C3d Testing and IgG Subclass Analysis of Human Leukocyte Antigen Antibodies. *Transpl Int* (2018) 31:1189–97. doi:10.1111/tri.13327
12. Roelen DL, DoxiadisII, Claas FHJ. Detection and Clinical Relevance of Donor Specific HLA Antibodies: a Matter of Debate. *Transpl Int* (2012) 25:604–10. doi:10.1111/j.1432-2277.2012.01491.x
13. Reindl-Schwaighofer R, Heinzel A, Gualdoni GA, Mesnard L, Claas FHJ, Oberbauer R. Novel Insights into Non-HLA Alloimmunity in Kidney Transplantation. *Transpl Int* (2020) 33:5–17. doi:10.1111/tri.13546
14. Patel R, Terasaki PI. Significance of the Positive Crossmatch Test in Kidney Transplantation. *N Engl J Med* (1969) 280:735–9. doi:10.1056/nejm196904032801401
15. Solez Ki. m., Axelsen RA, Benediktsson H, Burdick JF, Cohen AH, Colvin RB, et al. International Standardization of Criteria for the Histologic Diagnosis of Renal Allograft Rejection: the Banff Working Classification of Kidney Transplant Pathology. *Kidney Int* (1993) 44:411–22. doi:10.1038/ki.1993.259
16. Regele H, Exner M, Watschinger B, Wenter C, Wahrmann M, Österreicher C, et al. Endothelial C4d Deposition Is Associated with Inferior Kidney Allograft Outcome Independently of Cellular Rejection. *Nephrol Dial Transplant* (2001) 16:2058–66. doi:10.1093/ndt/16.10.2058
17. Racusen LC, Colvin RB, Solez K, Mihatsch MJ, Halloran PF, Campbell PM, et al. Antibody-Mediated Rejection Criteria - an Addition to the Banff '97 Classification of Renal Allograft Rejection. *Am J Transplant* (2003) 3:708–14. doi:10.1034/j.1600-6143.2003.00072.x
18. Solez K, Colvin RB, Racusen LC, Haas M, Sis B, Mengel M, et al. Banff 07 Classification of Renal Allograft Pathology: Updates and Future Directions. *Am J Transplant* (2008) 8:753–60. doi:10.1111/j.1600-6143.2008.02159.x
19. Haas M, Loupy A, Lefaucheur C, Roufosse C, Glotz D, Seron D, et al. The Banff 2017 Kidney Meeting Report: Revised Diagnostic Criteria for Chronic Active T Cell-Mediated Rejection, Antibody-mediated Rejection, and Prospects for Integrative Endpoints for Next-generation Clinical Trials. *Am J Transplant* (2018) 18:293–307. doi:10.1111/ajt.14625
20. Sis B, Mengel M, Haas M, Colvin RB, Halloran PF, Racusen LC, et al. Banff '09 Meeting Report: Antibody Mediated Graft Deterioration and Implementation of Banff Working Groups. *Am J Transplant* (2010) 10: 464–71. doi:10.1111/j.1600-6143.2009.02987.x
21. Haas M, Sis B, Racusen LC, Solez K, Glotz D, Colvin RB, et al. Banff 2013 Meeting Report: Inclusion of C4d-Negative Antibody-Mediated Rejection and Antibody-Associated Arterial Lesions. *Am J Transplant* (2014) 14: 272–83. doi:10.1111/ajt.12590
22. Loupy A, Haas M, Roufosse C, Naesens M, Adam B, Afrouzian M, et al. The Banff 2019 Kidney Meeting Report (I): Updates on and Clarification of Criteria for T Cell- and Antibody-mediated Rejection. *Am J Transplant* (2020) 20:2318–31. doi:10.1111/ajt.15898
23. Roufosse C, Simmonds N, Clahsen-van Groningen M, Haas M, Henriksen KJ, Horsfield C, et al. A 2018 Reference Guide to the Banff Classification of Renal Allograft Pathology. *Transplantation* (2018) 102:1795–814. doi:10.1097/tp.0000000000002366
24. Smith B, Cornell LD, Smith M, Cortese C, Geiger X, Alexander MP, et al. A Method to Reduce Variability in Scoring Antibody-Mediated Rejection in Renal Allografts: Implications for Clinical Trials - a Retrospective Study. *Transpl Int* (2018) 32:173–83. doi:10.1111/tri.13340
25. Preka E, Sekar T, Lopez Garcia SC, Shaw O, Kessaris N, Mamode N, et al. Outcomes of Paediatric Kidney Transplant Recipients Using the Updated 2013/2017 Banff Histopathological Classification for Antibody-Mediated Rejection. *Pediatr Nephrol* (2021) 36:2575–85. doi:10.1007/s00467-021-05103-x
26. De Serres SA, Noël R, Côté I, Lapointe I, Wagner E, Riopel J, et al. Banff Criteria for Chronic Active Antibody-Mediated Rejection: Assessment in a Real-Life Setting. *Am J Transplant* (2013) 16:1516–25. doi:10.1111/ajt.13624
27. De Kort H, Moran L, Roufosse C. The Role of Electron Microscopy in Renal Allograft Biopsy Evaluation. *Curr Opin Organ Transpl* (2015) 20:333–42. doi:10.1097/mot.0000000000000183
28. Schinstock CA, Mannon RB, Budde K, Chong AS, Haas M, Knechtle S, et al. Recommended Treatment for Antibody-Mediated Rejection after Kidney Transplantation: The 2019 Expert Consensus from the Transplantion Society Working Group. *Transplantation* (2020) 104:911–22. doi:10.1097/tp.0000000000003095
29. Solar-Cafaggi D, Marino L, Uribe-Uribe N, Morales-Buenrostro LE. Antibody-mediated Rejection in the Banff Classifications of 2007 and 2017: a Comparison of Renal Graft Loss Prediction Capability. *Transpl Immunol* (2018) 51:40–4. doi:10.1016/j.trim.2018.08.008
30. Sai K, OmotoShimizu KT, Shimizu T, Honda K, Tanabe K. The Impact of C4d-Negative Acute Antibody-Mediated Rejection on Short-Term Prognosis Among Kidney Transplant Recipients. *Nephrology* (2015) 20(Suppl. 2):16–9. doi:10.1111/nep.12473
31. Orandi BJ, Chow EHK, Hsu A, Gupta N, Van Arendonk KJ, Garonzik-Wang JM, et al. Quantifying Renal Allograft Loss Following Early Antibody-Mediated Rejection. *Am J Transplant* (2015) 15:489–98. doi:10.1111/ajt.12982
32. Orandi BJ, Alachkar N, Kraus ES, Naqvi F, Lonze BE, Lees L, et al. Presentation and Outcomes of C4d-Negative Antibody-Mediated Rejection after Kidney Transplantation. *Am J Transplant* (2016) 16: 213–20. doi:10.1111/ajt.13434
33. Everly MJ, Everly JJ, Arend LJ, Brailey P, Susskind B, Govil A, et al. Reducing De Novo Donor-specific Antibody Levels during Acute Rejection Diminishes Renal Allograft Loss. *Am J Transplant* (2009) 9:1063–71. doi:10.1111/j.1600-6143.2009.02577.x
34. Loupy A, Suberbielle-Boissel C, Hill GS, Lefaucheur C, Anglicheau D, Zuber J, et al. Outcome of Subclinical Antibody-Mediated Rejection in Kidney Transplant Recipients with Preformed Donor-specific Antibodies. *Am J Transplant* (2009) 9:2561–70. doi:10.1111/j.1600-6143.2009.02813.x
35. Wu K, Budde K, Schmidt D, Neumayer H-H, Rudolph B. The Relationship of the Severity and Category of Acute Rejection with Intimal Arteritis Defined in Banff Classification to Clinical Outcomes. *Transplantation* (2015) 99: e105–e114. doi:10.1097/tp.0000000000000640
36. Loupy A, Vernerey D, Tinel C, Aubert O, Duong van Huyen J-P, Rabant M, et al. Subclinical Rejection Phenotypes at 1 Year post-transplant and Outcome of Kidney Allografts. *J Am Soc Nephrol* (2015) 26:1721–31. doi:10.1681/asn.2014040399
37. Lesage J, Noël R, Lapointe I, Côté I, Wagner E, Désy O, et al. Donor-specific Antibodies, C4d and Their Relationship with the Prognosis of Transplant Glomerulopathy. *Transplantation* (2015) 99:69–76. doi:10.1097/tp.0000000000000310
38. Wavamunno MD, O'Connell PJ, Vitalone M, Fung CL-S, Allen RDM, Chapman JR, et al. Transplant Glomerulopathy: Ultrastructural Abnormalities Occur Early in Longitudinal Analysis of Protocol Biopsies. *Am J Transplant* (2007) 7:2757–68. doi:10.1111/j.1600-6143.2007.01995.x

39. Perkowska-Ptasińska A, Ciszek M, Chmura A, Galazka Z, Paczek L, Durlik M, et al. Transplant Glomerulopathy: Clinical and Pathological Correlations. *Transplant Proc* (2009) 41:141–9. doi:10.1016/j.transproceed.2008.10.052
40. Shimizu T, Ishida H, Toki D, Nozaki T, Omoto K, Tanabe K, et al. Clinical and Pathological Analyses of Transplant Glomerulopathy Cases. *Nephrology* (2014) 19:21–6. doi:10.1111/nep.12243
41. Hayde N, Bao Y, Pullman J, Ye B, Calder RB, Chung M, et al. The Clinical and Genomic Significance of Donor-specific Antibody-positive/C4d-Negative and Donor-specific Antibody-negative/C4d-Negative Transplant Glomerulopathy. *Clin J Am Soc Nephrol* (2013) 8:2141–8. doi:10.2215/cjn.04240413
42. Péfaur J, Díaz P, Panace R, Salinas P, Fiabane A, Quinteros N, et al. Early and Late Humoral Rejection: a Clinicopathologic Entity in Two Times. *Transplant Proc* (2008) 40:3229–36. doi:10.1016/j.transproceed.2008.03.123
43. Shimizu T, Ishida H, Shirakawa H, Omoto K, Tsunoyama K, Iida S, et al. Clinicopathological Analysis of Transplant Glomerulopathy Cases. *Clin Transplant* (2009) 23:39–43. doi:10.1111/j.1399-0012.2009.01008.x
44. John R, Konvalinka A, Tobar A, Kim SJ, Reich HN, Herzenberg AM. Determinants of Long-Term Graft Outcome in Transplant Glomerulopathy. *Transplantation* (2010) 90:757–64. doi:10.1097/tp.0b013e3181efcffd
45. Nair R, Fraer M, Agrawal N, Suneja M. Acute Transplant Glomerulopathy Is Associated with Antibody-Mediated Rejection and Poor Graft Outcome. *Transplant Proc* (2010) 42:3507–12. doi:10.1016/j.transproceed.2010.06.020
46. López Jiménez V, Fuentes L, Jiménez T, León M, Garcia I, Sola E, et al. Transplant Glomerulopathy: Clinical Course and Factors Relating to Graft Survival. *Transplant Proc* (2012) 44:2599–600. doi:10.1016/j.transproceed.2012.09.068
47. Kamal L, Broin PÓ, Bao Y, Ajaimy M, Lubetzky M, Gupta A, et al. Clinical, Histological, and Molecular Markers Associated with Allograft Loss in Transplant Glomerulopathy Patients. *Transplantation* (2015) 99:1912–8. doi:10.1097/tp.0000000000000598
48. Dobi D, Bodó Z, Kemény É, Bidiga L, Hódi Z, Szenohradszky P, et al. Peritubular Capillary Basement Membrane Multilayering in Early and Advanced Transplant Glomerulopathy: Quantitative Parameters and Diagnostic Aspects. *Virchows Arch* (2016) 469:563–73. doi:10.1007/s00428-016-2010-1
49. Halloran PF, Merino Lopez M, Barreto Pereira A. Identifying Subphenotypes of Antibody-Mediated Rejection in Kidney Transplants. *Am J Transplant* (2016) 16:908–20. doi:10.1111/ajt.13551
50. Toki D, Inui M, Ishida H, Okumi M, Shimizu T, Shirakawa H, et al. Interstitial Fibrosis Is the Critical Determinant of Impaired Renal Function in Transplant Glomerulopathy. *Nephrology* (2016) 21(Suppl. 1):20–5. doi:10.1111/nep.12765
51. Courant M, Visentin J, Linares G, Dubois V, Lepreux S, Guidicelli G, et al. The Disappointing Contribution of Anti-human Leukocyte Antigen Donor-specific Antibodies Characteristics for Predicting Allograft Loss. *Nephrol Dial Transplant* (2018) 33:1853–63. doi:10.1093/ndt/gfy088
52. Lubetzky M, Hayde N, Ó Broin P, Ajaimy M, Bao Y, Mohammed O, et al. Molecular Signatures and Clinical Outcomes of Transplant Glomerulopathy Stratified by Microvascular Inflammation and Donor-specific Antibody. *Clin Transplant* (2018) 33:e13469. doi:10.1111/ctr.13469
53. Sablik KA, Clahsen-Van Groningen MC, Looman CWN, Damman J, Roelen DL, van Agteren M, et al. Chronic-active Antibody-Mediated Rejection with or without Donor-specific Antibodies Has Similar Histomorphology and Clinical Outcome - a Retrospective Study. *Transpl Int* (2018) 31:900–8. doi:10.1111/tri.13154
54. Aubert O, Higgins S, Bouatou Y, Yoo D, Raynaud M, Viglietti D, et al. Archetype Analysis Identifies Distinct Profiles in Renal Transplant Recipients with Transplant Glomerulopathy Associated with Allograft Survival. *J Am Soc Nephrol* (2019) 30:625–39. doi:10.1681/asn.2018070777
55. Mengel M, Sis B, Haas M, Colvin RB, Halloran PF, Racusen LC, et al. Banff 2011 Meeting Report: New Concepts in Antibody-Mediated Rejection. *Am J Transplant* (2012) 12:563–70. doi:10.1111/j.1600-6143.2011.03926.x
56. Senev A, Van Loon E, Lerut E, Callemeyn J, Coemans M, Van Sandt V, et al. Risk Factors, Histopathological Features, and Graft Outcome of Transplant Glomerulopathy in the Absence of Donor-specific HLA Antibodies. *Kidney Int* (2021) 100:401–14. doi:10.1016/j.kint.2021.01.029
57. Zhang Q, Rudolph B, Choi M, Bachmann F, Schmidt D, Duerr M, et al. The Relationship between Proteinuria and Allograft Survival in Patients with Transplant Glomerulopathy: a Retrospective Single-center Cohort Study. *Transpl Int* (2021) 34:259–71. doi:10.1111/tri.13787
58. Kovács G, Devercelli G, Zelei T, Hirji I, Vokó Z, Keown PA. Association between Transplant Glomerulopathy and Graft Outcomes Following Kidney Transplantation: A Meta-Analysis. *PLoS One* (2020) 15:e0231646. doi:10.1371/journal.pone.0231646
59. Senev A, Coemans M, Lerut E, Van Sandt V, Daniëls L, Kuypers D, et al. Histological Picture of Antibody-mediated Rejection without Donor-specific anti-HLA Antibodies: Clinical Presentation and Implications for Outcome. *Am J Transplant* (2019) 19:763–80. doi:10.1111/ajt.15074
60. Mauiyyedi S, Crespo M, Collins AB, Schneeberger EE, Pascual MA, Saidman SL, et al. Acute Humoral Rejection in Kidney Transplantation: II. Morphology, Immunopathology, and Pathologic Classification. *J Am Soc Nephrol* (2002) 13:779–87. doi:10.1681/asn.v133779
61. Haas M, Ratner LE, Montgomery RA. C4d Staining of Perioperative Renal Transplant Biopsies1. *Transplantation* (2002) 74:711–7. doi:10.1097/00007890-200209150-00021
62. Djamali A, Muth BL, Ellis TM, Mohamed M, Fernandez LA, Miller KM, et al. Increased C4d in post-reperfusion Biopsies and Increased Donor Specific Antibodies at One-Week post Transplant Are Risk Factors for Acute Rejection in Mild to Moderately Sensitized Kidney Transplant Recipients. *Kidney Int* (2013) 83:1185–92. doi:10.1038/ki.2013.44
63. Kikić Ž, Kainz A, Kozakowski N, Oberbauer R, Regele H, Bond G, et al. Capillary C4d and Kidney Allograft Outcome in Relation to Morphologic Lesions Suggestive of Antibody-Mediated Rejection. *Clin J Am Soc Nephrol* (2015) 10:1435–43. doi:10.2215/CJN.09901014
64. Nickeleit V, Zeiler M, Gudat F, Thiel G, Mihatsch MJ. Detection of the Complement Degradation Product C4d in Renal Allografts: Diagnostic and Therapeutic Implications. *J Am Soc Nephrol* (2002) 13:242–51. doi:10.1681/asn.v131242
65. Dickenmann M, Steiger J, Descœudres B, Mihatsch M, Nickeleit V. The Fate of C4d Positive Kidney Allografts Lacking Histological Signs of Acute Rejection. *Clin Nephrol* (2006) 65:173–9. doi:10.5414/cnp65173
66. Dominy KM, Willicombe M, Al Johani T, Beckwith H, Goodall D, Brookes P, et al. Molecular Assessment of C4d-Positive Renal Transplant Biopsies without Evidence of Rejection. *Kidney Int Rep* (2019) 4:148–58. doi:10.1016/j.ekir.2018.09.005
67. Lefaucheur C, Loupy A, Vernerey D, Duong-Van-Huyen J-P, Suberbielle C, Anglicheau D, et al. Antibody-mediated Vascular Rejection of Kidney Allografts: a Population-Based Study. *Lancet* (2013) 381:313–9. doi:10.1016/s0140-6736(12)61265-3
68. Loupy A, Haas M, Solez K, Racusen L, Glotz D, Seron D, et al. The Banff 2015 Kidney Meeting Report: Current Challenges in Rejection Classification and Prospects for Adopting Molecular Pathology. *Am J Transplant* (2017) 17:28–41. doi:10.1111/ajt.14107
69. Le Quintrec M, Lionet A, Kamar N, Karras A, Barbier S, Buchler M, et al. Complement Mutation-AssociatedDe NovoThrombotic Microangiopathy Following Kidney Transplantation. *Am J Transplant* (20082008) 8:1694–701. doi:10.1111/j.1600-6143.2008.02297.x
70. Satoskar AA, Pelletier R, Adams P, Nadasdy GM, Brodsky S, Pesavento T, et al. De Novo thrombotic Microangiopathy in Renal Allograft Biopsies-Role of Antibody-Mediated Rejection. *Am J Transplant* (2010) 10:1804–11. doi:10.1111/j.1600-6143.2010.03178.x
71. Chua JS, Baelde HJ, Zandbergen M, Wilhelmus S, van Es LA, de Fijter JW, et al. Complement Factor C4d Is a Common Denominator in Thrombotic Microangiopathy. *J Am Soc Nephrol* (2015) 26:2239–47. doi:10.1681/asn.2014050429
72. Haas M, Mirocha J. Early Ultrastructural Changes in Renal Allografts: Correlation with Antibody-Mediated Rejection and Transplant Glomerulopathy. *Am J Transplant* (2011) 11:2123–31. doi:10.1111/j.1600-6143.2011.03647.x
73. Bagnasco SM, Zachary AA, Racusen LC, Arend LJ, Carter-Monroe N, Alachkar N, et al. Time Course of Pathologic Changes in Kidney Allografts of Positive Crossmatch HLA-Incompatible Transplant Recipients. *Transplantation* (2014) 97:440–5. doi:10.1097/01.tp.0000437177.40551.f4

74. Einecke G, Sis B, Reeve J, Mengel M, Campbell PM, Hidalgo LG, et al. Antibody-mediated Microcirculation Injury Is the Major Cause of Late Kidney Transplant Failure. *Am J Transplant* (2009) 9:2520–31. doi:10.1111/j.1600-6143.2009.02799.x
75. Sis B, Jhangri GS, Riopel J, Chang J, de Freitas DG, Hidalgo L, et al. A New Diagnostic Algorithm for Antibody-Mediated Microcirculation Inflammation in Kidney Transplants. *Am J Transplant* (2012) 12:1168–79. doi:10.1111/j.1600-6143.2011.03931.x
76. Verghese P, Dunn T, Najafian B, Kim Y, Matas A. The Impact of C4d and Microvascular Inflammation before We Knew Them. *Clin Transplant* (2013) 27:388–96. doi:10.1111/ctr.12111
77. De Kort H, Willicombe M, Brookes P, Dominy KM, Santos-Nunez E, Galliford JW, et al. Microcirculation Inflammation Associates with Outcome in Renal Transplant Patients WithDe NovoDonor-specific Antibodies. *Am J Transplant* (2013) 13:485–92. doi:10.1111/j.1600-6143.2012.04325.x
78. Sapir-Pichhadze R, Curran SP, John R, Tricco AC, Uleryk E, Laupacis A, et al. A Systematic Review of the Role of C4d in the Diagnosis of Acute Antibody-Mediated Rejection. *Kidney Int* (2015) 87:182–94. doi:10.1038/ki.2014.166
79. Matas AJ, Fieberg A, Mannon RB, Leduc R, Grande J, Kasiske BL, et al. Long-term Follow-up of the DeKAF Cross-sectional Cohort Study. *Am J Transplant* (2019) 19:1432–43. doi:10.1111/ajt.15204
80. Gloor JM, Sethi S, Stegall MD, Park WD, Moore SB, DeGoey S, et al. Transplant Glomerulopathy: Subclinical Incidence and Association with Alloantibody. *Am J Transplant* (2007) 7:2124–32. doi:10.1111/j.1600-6143.2007.01895.x
81. Torres IB, Salcedo M, Moreso F, Sellarés J, Castellá E, Azancot MA, et al. Comparing Transplant Glomerulopathy in the Absence of C4d Deposition and Donor-specific Antibodies to Chronic Antibody-Mediated Rejection. *Clin Transplant* (2014) 28:1148–54. doi:10.1111/ctr.12433
82. Baid-Agrawal S, Farris AB, 3rd, Pascual M, Mauiyyedi S, Farrell ML, Tolkoff-Rubin N, et al. Overlapping Pathways to Transplant Glomerulopathy: Chronic Humoral Rejection, Hepatitis C Infection, and Thrombotic Microangiopathy. *Kidney Int* (2011) 80:879–85. doi:10.1038/ki.2011.194
83. Olson JL. Renal Disease Caused by Hypertension. In: JC Jennette, editor. *Heptinstall's Pathology of the Kidney*. 7th ed. Philadelphia, PA: Wolters Kluwer (2015).
84. Einecke G, Reeve J, Sis B, Mengel M, Hidalgo L, Famulski KS, et al. A Molecular Classifier for Predicting Future Graft Loss in Late Kidney Transplant Biopsies. *J Clin Invest* (2010) 120:1862–72. doi:10.1172/jci41789
85. Roufosse CA, Shore I, Moss J, Moran LB, Willicombe M, Galliford J, et al. Peritubular Capillary Basement Membrane Multilayering on Electron Microscopy. *Transplantation* (2012) 94:269–74. doi:10.1097/tp.0b013e31825774ab
86. De Kort H, Willicombe M, Brookes P, Moran LB, Santos-Nunez E, Galliford JW, et al. Peritubular Capillary Basement Membrane Multilayering in Renal Allograft Biopsies of Patients with De Novo Donor-specific Antibodies. *Transplantation* (2016) 100:889–97. doi:10.1097/tp.0000000000000908
87. Sijpkens YW, Doxiadis II, Van Kemenade FJ, Zwinderman AH, De Fijter JW, Claas FH, et al. Chronic Rejection with or without Transplant Vasculopathy. *Clin Transplant* (2003) 17:163–70. doi:10.1034/j.1399-0012.2003.00039.x
88. Loupy A, Lefaucheur C, Vernerey D, Chang J, Hidalgo LG, Beuscart T, et al. Molecular Microscope Strategy to Improve Risk Stratification in Early Antibody-Mediated Kidney Allograft Rejection. *J Am Soc Nephrol* (2014) 25:2267–77. doi:10.1681/asn.2013111149
89. Yazdani S, Callemeyn J, Gazut S, Lerut E, de Loor H, Wevers M, et al. Natural Killer Cell Infiltration Is Discriminative for Antibody-Mediated Rejection and Predicts Outcome after Kidney Transplantation. *Kidney Int* (2019) 95:188–98. doi:10.1016/j.kint.2018.08.027
90. Sis B, Jhangri GS, Bunnag S, Allanach K, Kaplan B, Halloran PF. Endothelial Gene Expression in Kidney Transplants with Alloantibody Indicates Antibody-Mediated Damage Despite Lack of C4d Staining. *Am J Transplant* (2009) 9:2312–23. doi:10.1111/j.1600-6143.2009.02761.x
91. Dominy KM, Roufosse C, De Kort H, Willicombe M, Brookes P, Behmoaras JV, et al. Use of Quantitative Real Time Polymerase Chain Reaction to Assess Gene Transcripts Associated with Antibody-Mediated Rejection of Kidney Transplants. *Transplantation* (2015) 99:1981–8. doi:10.1097/tp.0000000000000621
92. Halloran PF, Pereira AB, Chang J, Matas A, Picton M, De Freitas D, et al. Microarray Diagnosis of Antibody-Mediated Rejection in Kidney Transplant Biopsies: an International Prospective Study (INTERCOM). *Am J Transplant* (2013) 13:2865–74. doi:10.1111/ajt.12465
93. Reeve J, Böhmig GA, Eskandary F, Einecke G, Lefaucheur C, Loupy A, et al. Assessing Rejection-Related Disease in Kidney Transplant Biopsies Based on Archetypal Analysis of Molecular Phenotypes. *JCI Insight* (2017) 2:e94197. doi:10.1172/jci.insight.94197
94. Lerut E, Naesens M, Kuypers DR, Vanrenterghem Y, Van Damme B. Subclinical Peritubular Capillaritis at 3 Months Is Associated with Chronic Rejection at 1 Year. *Transplantation* (2007) 83:1416–22. doi:10.1097/01.tp.0000266676.10550.70
95. Haas M, Montgomery RA, Segev DL, Rahman MH, Racusen LC, Bagnasco SM, et al. Subclinical Acute Antibody-Mediated Rejection in Positive Crossmatch Renal Allografts. *Am J Transplant* (2007) 7:576–85. doi:10.1111/j.1600-6143.2006.01657.x
96. Loupy A, Hill GS, Suberbielle C, Charron D, Anglicheau D, Zuber J, et al. Significance of C4d Banff Scores in Early Protocol Biopsies of Kidney Transplant Recipients with Preformed Donor-specific Antibodies (DSA). *Am J Transplant* (2011) 11:56–65. doi:10.1111/j.1600-6143.2010.03364.x
97. Cosio FG, El Ters M, Cornell LD, Schinstock CA, Stegall MD. Changing Kidney Allograft Histology Early Posttransplant: Prognostic Implications of 1-Year Protocol Biopsies. *Am J Transplant* (2016) 16:194–203. doi:10.1111/ajt.13423
98. Papadimitriou JC, Drachenberg CB, Ramos E, Kukuruga D, Klassen DK, Ugarte R, et al. Antibody-Mediated Allograft Rejection. *Transplantation* (2013) 95:128–36. doi:10.1097/tp.0b013e3182777f28
99. Tsuji T, Yanai M, Itami H, Ishii Y, Akimoto M, Fukuzawa N, et al. Microvascular Inflammation in Early Protocol Biopsies of Renal Allografts in Cases of Chronic Active Antibody-Mediated Rejection. *Nephrology* (2015) 20(Suppl. 2):26–30. doi:10.1111/nep.12450
100. Parajuli S, Joachim E, Alagusundaramoorthy S, Blazel J, Aziz F, Garg N, et al. Subclinical Antibody-Mediated Rejection after Kidney Transplantation: Treatment Outcomes. *Transplantation* (2019) 103:1722–9. doi:10.1097/tp.0000000000002566
101. Wiebe C, Rush DN, Gibson IW, Pochinco D, Birk PE, Goldberg A, et al. Evidence for the Alloimmune Basis and Prognostic Significance of Borderline T Cell-Mediated Rejection. *Am J Transplant* (2020) 20:2499–508. doi:10.1111/ajt.15860
102. Burton SA, Amir N, Asbury A, Lange A, Hardinger KL. Treatment of Antibody-Mediated Rejection in Renal Transplant Patients: a Clinical Practice Survey. *Clin Transplant* (2015) 29:118–23. doi:10.1111/ctr.12491

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10139

# Allograft Function as Endpoint for Clinical Trials in Kidney Transplantation

*Luuk Hilbrands[1], Klemens Budde[2], Maria Irene Bellini[3], Fritz Diekmann[4], Lucrezia Furian[5], Josep Grinyó[6], Uwe Heemann[7], Dennis A. Hesselink[8], Alexandre Loupy[9], Rainer Oberbauer[10], Liset Pengel[11], Marlies Reinders[8], Stefan Schneeberger[12] and Maarten Naesens[13]**

[1]Department of Nephrology, Radboud University Medical Center, Nijmegen, Netherlands, [2]Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany, [3]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy, [4]Department of Nephrology and Kidney Transplantation, Vall d'Hebrón University Hospital, Barcelona, Spain, [5]Kidney and Pancreas Transplantation Unit, University of Padua, Padua, Italy, [6]Department of Clinical Sciences, University of Barcelona, Barcelona, Spain, [7]Department of Nephrology, Technical University of Munich, Munich, Germany, [8]Department of Internal Medicine, Erasmus MC Transplant Institute, University Medical Center Rotterdam, Rotterdam, Netherlands, [9]Paris Translational Research Center for Organ Transplantation, Hôpital Necker, Paris, France, [10]Department of Nephrology and Dialysis, Medical University of Vienna, Vienna, Austria, [11]Centre for Evidence in Transplantation, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom, [12]Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria, [13]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

Clinical study endpoints that assess the efficacy of interventions in patients with chronic renal insufficiency can be adopted for use in kidney transplantation trials, given the pathophysiological similarities between both conditions. Kidney dysfunction is reflected in the glomerular filtration rate (GFR), and although a predefined (e.g., 50%) reduction in GFR was recommended as an endpoint by the European Medicines Agency (EMA) in 2016, many other endpoints are also included in clinical trials. End-stage renal disease is strongly associated with a change in estimated (e)GFR, and eGFR trajectories or slopes are increasingly used as endpoints in clinical intervention trials in chronic kidney disease (CKD). Similar approaches could be considered for clinical trials in kidney transplantation, although several factors should be taken into account. The present Consensus Report was developed from documentation produced by the European Society for Organ Transplantation (ESOT) as part of a Broad Scientific Advice request that ESOT submitted to the EMA in 2020. This paper provides a contemporary discussion of primary endpoints used in clinical trials involving CKD, including proteinuria and albuminuria, and evaluates the validity of these concepts as endpoints for clinical trials in kidney transplantation.

**Keywords: kidney transplantation, graft function, graft dysfunction, clinical study, endpoints**

OPEN ACCESS

***Correspondence:**
*Maarten Naesens*
*maarten.naesens@kuleuven.be*

***Received:** 21 October 2021*
***Accepted:** 11 January 2022*
***Published:** 20 May 2022*

***Citation:***
*Hilbrands L, Budde K, Bellini MI, Diekmann F, Furian L, Grinyó J, Heemann U, Hesselink DA, Loupy A, Oberbauer R, Pengel L, Reinders M, Schneeberger S and Naesens M (2022) Allograft Function as Endpoint for Clinical Trials in Kidney Transplantation. Transpl Int 35:10139. doi: 10.3389/ti.2022.10139*

## INTRODUCTION

As with progressive chronic disease of native kidneys, chronic graft failure results in end-stage renal disease (ESRD) with the need for kidney replacement therapy in the form of dialysis or repeat transplantation. Pathological processes that characterize the late course of graft failure are loss of nephrons, glomerulosclerosis of the remaining nephrons, and interstitial fibrosis and tubular atrophy

(1). Essentially, these processes are no different between graft dysfunction and other forms of chronic kidney disease (CKD).

Loss of viable nephrons is reflected in a reduced glomerular filtration rate (GFR), therefore late kidney graft failure is inevitably preceded by a decline in GFR. Notably, the annual rate of eGFR decline in incident dialysis patients with graft failure is higher than in transplant-naïve incident dialysis patients (2): this is potentially related to the hypothesis that multiple factors — or more severe factors — contribute to nephron loss in transplanted patients compared with those who have chronic disease of native kidneys.

The present article provides an overview of primary endpoints used in clinical trials involving CKD, including a contemporary perspective on endpoints for assessing graft dysfunction after kidney transplantation. Biomarkers that have meaningful associations with graft failure are discussed.

## PRIMARY ENDPOINTS FOR SECONDARY PREVENTION OF CHRONIC RENAL INSUFFICIENCY

The guideline EMA/CHMP/500825/2016 (3) addresses the clinical development of compounds designed to prevent (or slow) processes implicated in chronic renal insufficiency. In its choice of endpoints, EMA distinguishes between primary and secondary prevention of chronic renal insufficiency. In kidney transplant recipients with CKD, the relevant objective is secondary prevention. The recommended primary endpoint in secondary prevention is the time to a predefined loss in GFR, such as a 50% loss. Other (lower) proportions might be used, provided the magnitude is qualified for a specific primary disease or patient population (e.g., extrapolating adult data to pediatric patients). Therefore, three endpoints for graft (dys)function from the EMA 2016 guideline are particularly relevant for kidney transplant recipients (3):

- Kidney function at different timepoints (e.g., 6, 12, and 24 months; 3 and 5 years)
- Proteinuria incidence or worsening
- Time to reach different CKD stages (representing progression of renal damage).

Notably, for primary prevention studies (defined as CKD prevention in patients without any sign of kidney damage), the European Medicines Agency (EMA) recommends using a clinically meaningful and stable GFR loss rate (measured either via slope or time-to-event analyses) as the primary endpoint (3). However, since the EMA guideline was released, additional literature has been published on the choice of CKD endpoints, and various endpoints have been used in clinical trials. There has been increasing use of the eGFR slope in secondary prevention trials, including studies evaluating graft function. These are discussed below.

## EVALUATING GRAFT FUNCTION

As initial decline in kidney function is asymptomatic, and clinical manifestations of renal insufficiency occur late in the disease course, general definitions of kidney disease focus on measures of function (e.g., GFR) or damage (e.g., proteinuria, morphological abnormalities). Several CKD biomarkers indicate levels of kidney damage (e.g., active urinary sediment, presence of proteinuria or albuminuria, leakage markers) or functional status (e.g., failure to filtrate plasma or endogenous substances, absorb primary urine, secrete hydrogen ions, or contribute to endocrine function) (3).

Filtration reflects the main function of the kidneys, and GFR is also used as an indicator of kidney function in grafts. Markers for calculating measured (m)GFR must be freely filtered in the glomeruli and not reabsorbed, secreted, or metabolized by renal cells. Although exogenous substances including inulin and iothalamate fulfill these criteria, their analysis requires intravenous infusion; methods to measure their concentrations are costly and are not universally available nor necessarily error free.

## GLOMERULAR FILTRATION RATE

### Creatinine and Cystatin C

Creatinine and cystatin C measurements are widely used to assess GFR in clinical or research settings, in every relevant patient population (including kidney transplant recipients). Creatinine is a breakdown product of creatine from muscle cells and is largely removed from the blood by glomerular filtration. Therefore, the serum creatinine level, which is easily measured, is a useful reflection of GFR and is traditionally analyzed as an indicator of kidney function. Several decades ago, cystatin C — released by all nucleated cells — was also shown to be a reasonable marker of kidney function. Importantly, however, neither creatinine nor cystatin C meet the requirements of an ideal filtration marker (4). Creatinine levels depend not only on GFR, but also on muscle mass and dietary meat intake; cystatin C levels can also increase with corticosteroid treatment, which is frequently administered after kidney transplantation (5).

Creatinine clearance over a 24-h period can be used as a surrogate marker of GFR. However, creatinine is also secreted, which leads to GFR overestimation. Moreover, 24-h urine collection is burdensome, and inaccuracies in collection cause discrepancies between creatinine clearance and GFR. To overcome such limitations, equations have been developed to calculate eGFR (3, 6, 7). Common denominators in these formulas are serum creatinine and/or cystatin C levels; additional factors used in different combinations are sex, weight, age, and ethnicity. The most frequently used equations derive from the Modification of Diet in Renal Disease (MDRD) study (6) or the CKD-EPI formula (7), which have been validated in transplanted populations (8, 9). However, difficulties associated with measurements based on creatinine and cystatin C levels translate into limitations when applying these formulas, resulting in suboptimal agreement between eGFR and mGFR in the individual patient. Importantly, when it is essential to determine GFR precisely (e.g., when an anticipated decline in function will occur slowly, in longitudinal studies, or when there is considerable variation in non-GFR determinants of biomarkers employed for

estimation), the EMA recommends mGFR rather than eGFR (3).

## eGFR Versus mGFR

In kidney transplantation populations, the use of formulas such as MDRD and CKD-EPI to calculate eGFR (8, 9) is hampered by post-transplant differences in body composition (caused by protein catabolic effects of corticosteroids or edema) and inhibition of tubular creatinine secretion by trimethoprim (which is frequently administered). Nevertheless, they are widely used to evaluate GFR in clinical trials of kidney transplantation.

For accurate assessment of kidney function at a given time point in an individual, mGFR is undoubtedly the best available method (10), but this is difficult to undertake in routine practice. However, for comparing cohorts in clinical trial settings, the precise value in the individual may not be required: average eGFR values perform as well in study group comparisons as average mGFR values (11–15). The EMA/CHMP proposes that mGFR is performed in a prespecified subset of patients to confirm eGFR, with creatinine-based eGFR used in preference to cystatin C-based estimations (as creatinine-based eGFR is better characterized). Regardless of the methodology used for eGFR, the influence of confounders on data interpretation should be considered (3).

When selecting an outcome measure of kidney function for clinical research, it is important to know the strength of the relationship between each measure and the occurrence of hard endpoints such as ESRD. No studies show that one-time determination of mGFR is more strongly associated with future ESRD than eGFR, and mGFR has potential limitations, such as the complexity of evaluating large trials. Another major drawback of mGFR evaluation is the impossibility of calculating slopes over time (see below), which requires many repeated measurements.

There is limited agreement between decline in mGFR and eGFR. The Chronic Renal Insufficiency Cohort (CRIC) study compared associations between longitudinal changes (two measurements in 24 months) in eGFR and mGFR (urinary iothalamate clearance) with ESRD risk (16). The strongest association was found for changes in eGFR, which may be explained by higher precision (i.e., less variability) in GFR measurement using eGFR compared with mGFR. In a study of octreotide long-acting release in patients with autosomal-dominant polycystic kidney disease, similar slopes were observed for mGFR and eGFR in intervention and control groups during the 3-year follow-up period (17); comparable studies are not available for the transplant population. In accordance with the prevailing view in the nephrological community, we consider the use of eGFR as a suitable alternative for the practically cumbersome and more costly mGFR, in longitudinal studies and for comparison of investigational groups.

# PROTEINURIA

Proteinuria is generally measured as the albumin or total protein concentration in a spot sample or in urine collected during a specified time period (e.g., 24 h); in the latter case, the excretion rate of albumin or protein can be calculated. Consequently, EMA/CHMP guidelines state that proteinuria should be assessed quantitatively, using a timed or untimed (spot) urine collection (3). When albumin or protein concentrations are measured in a spot sample, it is important to correct for the urine concentration by simultaneously measuring the creatinine concentration. Accordingly, measurements are expressed as the albumin:creatine ratio (ACR), or protein:creatine ratio (PCR). There is no reason to consider adopting a different policy in kidney transplant recipients.

Since collection of timed urine samples is inconvenient and error prone, use of spot samples has gained popularity. Studies in people with diabetes mellitus, immunoglobulin (Ig) A nephropathy, and a mixed cohort of patients with CKD show that measuring ACR in a morning spot sample is at least equal to measuring 24-h albumin or protein excretion for predicting CKD progression (18–20). In a cohort of 207 kidney transplant recipients, spot and 24-h measurements of albumin and protein excretion were similar predictors of doubling of serum creatinine level and graft loss (21). Therefore, spot sampling can also be recommended in kidney transplant recipients.

Generally, EMA/CHMP guidelines prefer ACR to PCR, especially at low levels of proteinuria, acknowledging that PCR may be the best way to characterize kidney injury (e.g., diabetic nephropathy). Timed urine collection and testing is required after any positive ACR/PCR result to confirm the findings, although repeat ACR/PCR could also be considered. EMA/CHMP guidelines also state that timed urine sample testing would be necessary to assess therapeutic efficacy during a clinical study (3).

# PRIMARY ENDPOINTS OF GRAFT FUNCTION IN RELATION TO CHRONIC KIDNEY DISEASE LITERATURE

As mentioned above, progressive decline in kidney graft function in many aspects resembles the course of dysfunction in native kidney disease. However, compared with CKD, it must be realized that after kidney transplantation the course of kidney function is more subject to acute events such as infection and rejection, as well as to changes in immunosuppressive therapy. Literature on kidney endpoints has largely focused on CKD, but data for transplant recipients are available. Data on kidney function endpoints post transplantation were extracted from 213 reports from randomized controlled trials (RCTs) published between 2010 and 2014, comparing immunosuppressive interventions (22). In 44 reports, a measure of kidney function (usually eGFR) was the primary endpoint, although some had other primary endpoints such as graft survival. **Table 1** summarizes RCTs in transplantation published after 2014 with kidney function as the primary endpoint (14, 15, 23–30).

## eGFR as the Endpoint

Doubling of the serum creatinine level is commonly an endpoint in clinical studies of kidney disease, including transplantation: it is considered analogous to a prespecified reduction in eGFR, and

**TABLE 1 |** RCTs in kidney transplantation with renal function as primary endpoint, published after 2014 (14, 15, 23–30).

| Study | Population | Intervention/control | Duration | Renal endpoint | Finding | Comments |
|---|---|---|---|---|---|---|
| APOLLO (23) | *N* = 93; >6 months after Tx sCr <2.5 mg/dl | I: conversion from CNI to EVR<br>C: continuation of CNI | 12 months | eGFR: Nankivell | NSD | Premature termination due to slow recruitment. Higher eGFR — MDRD in EVR group |
| CENTRAL (14) | *N* = 212; 7 weeks after Tx | I: conversion from CsA to EVR<br>C: continuation of CsA | 3 years | Change in measured GFR by iohexol or $^{51}$Cr-EDTA clearance from randomization to 36 months | NSD | High rate of study withdrawals. Benefit in renal function in EVR group in on-treatment analysis |
| SPIESSER(24) | *N* = 145; *dn*Tx | I: SRL<br>C: CsA | 12 months | eGFR: Nankivell | NSD | Benefit in renal function in EVR group in on-treatment analysis |
| Tedesco-Silva et al. (25) | *N* = 256; 90–150 days after Tx | I: conversion from TAC to SRL<br>C: continuation of TAC | 24 months | eGFR: MDRD change >5 ml/1.73 m$^2$ in on-therapy population (*n* = 195) | NSD | High discontinuation rate in SLR group |
| Knoll et al. (26) | *N* = 212; >3 months after Tx; eGFR ≥20 ml/min/1.73 m$^2$ and proteinuria ≥0.2 g/d | I: ramipril<br>C: placebo | 48 months | Composite endpoint: doubling of sCr, ESRD, or death | NSD | Small numbers per group |
| ELEVATE (27) | *N* = 715; 10–14 weeks after Tx | I: conversion from CNI to EVR<br>C: continuation of CNI | 24 months | Change in eGFR — MDRD from randomization to 12 m | NSD | Significantly higher eGFR in EVR group vs CsA subgroup |
| ADHERE (15) | *N* = 730; 28 days after Tx | I: TAC (8–12 ng/ml until Day 41 and then 6–10 ng/ml) + SRL<br>C: continuation of TAC (8–12 ng/ml) + MMF | 12 months | mGFR by iohexol clearance | NSD | High withdrawal rate in the intervention group |
| 3C STUDY (28) | *N* = 394; 6 m after Tx | I: conversion from TAC to SRL<br>C: continuation of TAC | 18 months | eGFR—MDRD | NSD | Significantly better renal function in SRL group in on-treatment analysis |
| BORTEJECT (29) | *N* = 44; presence of DSA and morphologic features of AMR ≥180 days after Tx, eGFR >20 ml/min/1.73 m$^2$ | I: bortezomib<br>C: placebo | 24 months | Slope of eGFR—Mayo equation | NSD | Small sample size |
| TRANSFORM (30) | *N* = 2037; *dn*Tx | I: EVR + reduced-dose CNI<br>C: MPA + standard-dose CNI | 24 months | Composite of treated BPAR or eGFR—MDRD <50 ml/min/1.73 m$^2$ at 12 months | NSD | No difference in eGFR |

*AMR, antibody-mediated rejection; C, control group; CNI, calcineurin inhibitor; CsA, cyclosporine; dn,* de novo*; DSA, donor-specific antibodies; e, estimated; ESRD, end-stage renal disease; EVR, everolimus; GFR, estimated glomerular filtration rate; I, intervention group; m, measured; MDRD, Modification of Diet in Renal Disease; MPA, mycophenolic acid; NSD, no statistical difference; sCR, serum creatinine; SRL, sirolimus; TAC, tacrolimus; Tx, transplantation.*

is often part of a composite outcome together with initiation of kidney replacement therapy and death from a renal cause (31, 32). However, doubling of serum creatinine is a late event in the progression of kidney insufficiency, and using this measure requires lengthy follow-up and/or very large numbers of patients. Alternative endpoints include the use of less steep (e.g., 30%) declines in eGFR, which have been strongly associated with ESRD (33, 34).

Consistent with findings observed in CKD (33), others have demonstrated that graft failure can be predicted not only by eGFR level at a given time point, but also by decline in eGFR within a relatively short period. For example, data from the Australian and New Zealand Dialysis and Transplant Registry (7,949 transplants) indicated that a ≥30% decline in eGFR between years 1 and 3 post-transplantation was strongly associated with subsequent death-censored graft failure and mortality (35).

The Clinical Trials in Organ Transplantation consortium showed that a 20–40% decline in eGFR at 3–24 or 6–24 months post transplantation was significantly associated with graft loss at 2–5 years, and with absolute eGFR at 5 years (36). The relationship between changes in eGFR and graft loss at 5 years was confirmed in the Genomics of Chronic Renal Allograft Rejection study (36). This suggests that the recommendation of a ≥30% eGFR decline over 2 years as being an acceptable outcome measure in CKD trials could be extended to studies involving kidney transplant recipients.

In some circumstances eGFR may not be a valid surrogate endpoint, although potential solutions for some situations have been formulated (34). For example, in kidney transplantation studies, eGFR decline may not be an ideal surrogate endpoint when drugs that affect muscle mass are administered, such as when muscle atrophy results from corticosteroid treatment (37). In addition, commencing or

discontinuing angiotensin-converting enzyme inhibitors, angiotensin receptor blockers (ARBs), or calcineurin inhibitors can acutely affect GFR (38), which could have implications for trial design when eGFR is a surrogate endpoint: inclusion of a run-in period may be warranted.

Several RCTs involving kidney transplant recipients have used eGFR, or a change in eGFR, as a single primary endpoint (23–25, 27, 29, 30, 39). In most cases, eGFR was calculated using the MDRD or Nankivell formulas. Studies compared different immunosuppressive regimens, either *de novo* or starting at a specific time after transplantation. The fact that a significant difference in the primary outcome measure was absent in all but one study is probably not a weakness of the chosen endpoint, but rather illustrates that current immunosuppressive regimens are generally equivalent with respect to short-term graft function.

Therefore, in most circumstances, post-transplantation eGFR seems to be an appropriate endpoint for evaluating graft dysfunction. A systematic review concluded that post-transplantation eGFR (at 12 months) is associated with risk for overall or death-censored graft loss, and all-cause mortality, in univariate and multivariate analyses (40), although such a highly significant association does not necessarily translate into good predictive capability (41). The magnitude of the association between reduced GFR and outcomes was greater for death-censored graft loss versus overall loss, and for graft loss compared with overall patient mortality (40).

## eGFR Trajectories as the Endpoint

Clinical studies in diabetic nephropathy, hypertension and CKD, and polycystic kidney disease use the eGFR slope to evaluate the efficacy of interventions that aim to slow progression of kidney insufficiency (32, 42–45).

In 2017, the KDIGO conference *Challenges in the Conduct of Clinical Trials in Nephrology* concluded that change in eGFR over time was a practical and acceptable method for assessing kidney disease progression (46). KDIGO considered CKD stage and progression rate as determinants of the most useful outcome measure (46). When there is a markedly reduced kidney function (eGFR <45 ml/min/1.73 $m^2$) and/or a rapid decline in GFR (>5 ml/min/1.73 $m^2$ per year), a composite endpoint consisting of a 30–40% decline in eGFR or the occurrence of ESRD failure is a robust and feasible outcome. When there is a slow decline in kidney function, using the GFR slope as the outcome measure may circumvent the need for lengthy follow-up and/or recruitment of very large numbers of patients. The same considerations can be applied in kidney transplantation studies.

In March 2018, several meta-analyses based on individual patient data were conducted in preparation for the workshop *Change in Albuminuria and GFR as End Points for Clinical Trials in Early Stages of CKD*, which evaluated surrogate endpoints for trials of CKD progression (47, 48). One meta-analysis (14 CKD cohorts) showed that, compared with a rapid decline in eGFR, a slower decline is associated with a lower risk of subsequent ESRD, even in participants with eGFR ≥60 ml/min/1.73 $m^2$ (49). A second meta-analysis of 47 RCTs evaluated the GFR slope as a surrogate endpoint for trials examining effects on CKD progression (50). This showed, with sufficiently large sample sizes, that treatment effects on the GFR slope from baseline and 3-month follow-up of 0.5–1.0 ml/min/1.73 $m^2$ per year strongly predicted benefits on clinical endpoints such as doubling of serum creatinine, GFR <15 ml/min/1.73 $m^2$, or ESRD. Using statistical simulations of GFR trajectories based on data from the 47 RCTs, the GFR slope performed better than clinical endpoints when patients' initial GFRs were high and not acutely affected by treatment (51). Although cohorts that formed the basis of these meta-analyses did not include kidney transplant recipients, there was a large variation in underlying causes of CKD. This makes it reasonable to assume that the study conclusions would apply to patients with graft dysfunction as a particular form of CKD.

A theoretical advantage of using the eGFR slope as the endpoint, rather than a time-to-event endpoint (e.g., ESRD), is that the decision to initiate dialysis or (re)transplantation can be affected by factors other than GFR. The effect of an intervention on the eGFR slope may therefore better reflect the true effect on kidney graft function.

Importantly, suitability of the eGFR slope as surrogate endpoint depends on patterns of acute and chronic phases of the slope, in the context of the specific disease and potential pharmacodynamic impact of the investigational compound on the slope. The total eGFR slope reflects the slope from time of randomization, i.e., across the entire study period; the chronic slope calculation starts later and is less affected by acute changes in eGFR during the initial phase post randomization. In this respect, transplant recipients likely differ from patients with native CKD. The chance of a non-linear decline in eGFR is probably higher in kidney transplant recipients as a result of acute events such as infection, rejection and initiation or withdrawal of drugs that have acute effects on kidney function, including immunosuppressive agents. This can also impact the number of eGFR measurements needed to calculate the eGFR slope. Such information should be available to evaluate the usefulness of eGFR slope after kidney transplantation and to judge the validity of the chronic slope, versus the total slope. The EMA has indicated that the total slope is generally favored over the chronic slope, because the total slope minimizes possible biases introduced when post-randomization events (e.g., death) or acute changes in eGFR on investigational drug initiation are not considered (48).

The only study in kidney transplant recipients to use the eGFR slope as the primary outcome measure was a clinical trial of bortezomib in late antibody-mediated rejection (AMR) (29). In preparation for a placebo-controlled trial investigating clazakizumab as a treatment for chronic AMR, a data-modeling exercise evaluated the relationship between rate of eGFR decline and risk of graft failure. This was a historical prospective cohort study investigating the relationship between change in eGFR (estimated using the MDRD 4 equation) and risk of graft failure in kidney transplant recipients diagnosed with acute/active (a)AMR (52). The primary analysis used data from 91 patients with biopsy-proven aAMR and baseline eGFR ≥25 ml/min/1.73 $m^2$, with a minimum of 3 years' follow-up data. Both a linear mixed-effects model to describe eGFR decline and a joint model, involving longitudinal eGFR and its

rate of decline, were constructed. The joint model predicted that the baseline eGFR and its rate of decline (slope change per month) following an aAMR diagnosis significantly predicts risk of both death-censored and all-cause graft failure. Using the modeling results for all-cause graft loss, the mean eGFR decline from baseline to month 12 after AMR diagnosis was 0.75 ml/min/1.73 $m^2$ per month. Using these data, and assuming a 50% reduction in the rate of eGFR decline with clazakizumab, a sample size was calculated for an interim analysis of the 52-week eGFR endpoint. This is an example of how modern statistical techniques can optimize study design in a specific patient population. For these techniques to be used, data are required on the natural course (i.e., without intervention) of kidney function in the population of interest, which should be sufficiently large to accurately define the natural disease course.

## mGFR as a Primary Endpoint

Few RCTs in kidney transplantation have used a change in mGFR as the primary endpoint; mGFR assessment was based on iohexol or $^{51}$Cr-EDTA clearance (14, 15). Although eGFR values were markedly higher than mGFR values in both studies, the conclusions were not affected when eGFR was used instead of mGFR; thus, no advantage for using mGFR was demonstrated. The BENEFIT and BENEFIT-EXT studies (11, 12) and the Spare The Nephron trial (13) also used mGFR as primary endpoint; the latter illustrated the difficulty in obtaining mGFR data, as no values were available in nine of 112 (8.0%) and 25/116 (21.6%) patients in the mycophenolate mofetil (MMF)/sirolimus and MMF/calcineurin inhibitor (CNI) groups (13). In the BENEFIT study, missing mGFR values were imputed using eGFR values, although the exact magnitude of this imputation is not available (11). Measured GFR is no longer used as a kidney function endpoint in kidney transplantation studies because of limitations mentioned earlier and availability of different methods. In summary, ESOT considers eGFR as the most useful marker for post-transplantation kidney function.

## Proteinuria as the Endpoint

While proteinuria can be considered as a surrogate marker for severity of glomerular damage, proteinuria can also directly contribute to kidney injury and decline in kidney function (53). Although not formally proven, this finding probably also holds true for kidney transplantation populations.

In a large cohort (31,372 individuals from a general population; two or more ACR measurements in 2 years), a fourfold increase in ACR was associated with a threefold heightened risk of ESRD during a median 3 years of follow-up (54). A reduction in proteinuria is also known to protect patients with various forms of renal disease from kidney function decline (55).

In IgA nephropathy, proteinuria is the most widely recognized risk factor for progression to ESRD. Analysis of 13 controlled intervention trials in IgA nephropathy showed an association between treatment effects on percentage reduction of proteinuria, and on a composite of time to doubling of serum creatinine, ESRD, or death (56). Similarly, a meta-analysis of 41 randomized trials in CKD supported use of change in albuminuria level as a surrogate endpoint for CKD progression, particularly in patients with high baseline albuminuria (57). A European Regulator's perspective on the potential of change in albuminuria as endpoint for clinical trials in CKD has been published (48).

Unlike specific diseases in native kidneys, causes of proteinuria after kidney transplantation are diverse. During the early months after transplantation there may be some contribution from proteinuric native kidneys, but major causes of proteinuria are chronic rejection, recurrence of proteinuric disease, or *de novo* glomerulopathy. Nevertheless, an association between proteinuria and progression to ESRD (demonstrated particularly in diabetes mellitus and IgA nephropathy) has been observed in several cohorts of kidney transplant recipients (58–60). Such findings were confirmed in a *post hoc* analysis of the FAVORIT trial (3,511 participants followed over 4 years), which found that an elevated baseline ACR is independently associated with graft failure, cardiovascular disease, and death (61).

In contrast to studies investigating chronic disease in native kidneys, no studies in kidney transplantation have demonstrated a beneficial effect of proteinuria reduction on progression to ESRD. A clinical trial of ramipril versus placebo in 213 kidney transplant recipients with (mean proteinuria ≥0.2 g/day) showed no difference in the primary outcome (a composite of doubling of serum creatinine, ESRD or death), despite some reduction in mean proteinuria (26). Evidential support for using proteinuria as a post–kidney transplantation endpoint is weaker than evidence for its use in studies of CKD in native kidneys. More evidence is needed from larger cohorts before proteinuria could be proposed for use in kidney transplantation clinical trials.

## Combined eGFR and Proteinuria Endpoint

The KDIGO 2012 guidelines updated the classification system for CKD to include albuminuria, stating that, for the general population, risk of adverse outcomes (mortality, progression to ESRD) at a given eGFR increases with higher levels of albuminuria.

Although studies indicate its promise (26, 61), the combination of eGFR and proteinuria (either as absolute values or as changes from baseline) has not been used as an endpoint in kidney disease clinical trials. Data from the ADVANCE study showed that, in patients with type 2 diabetes mellitus, the 2-year change in eGFR and ACR more strongly predicted the risk of ESRD during a median follow-up of 7.7 years than either of these changes alone (62). A limitation of this study is that the combination of worsening of eGFR and increase in urinary ACR, as well as major kidney events, occurred in only 1% of the study population. Additional studies are required to determine when the combination of changes in proteinuria and eGFR can be used as a surrogate outcome in a broad spectrum of kidney diseases.

The interaction between eGFR and proteinuria — as demonstrated in participants with diabetes in the ADVANCE study — was also observed in kidney transplant recipients. An analysis of linked databases in Canada ($N = 900$) found that rates of death-censored graft loss also increased with lower levels of kidney function at 1 year after transplantation (63). Moreover, within each eGFR category, adjusted rates increased with higher

levels of proteinuria. Risk of death-censored graft loss was 49-fold higher for kidney transplant recipients with an eGFR of 15–29 ml/min/1.73 $m^2$ and severely increased albuminuria, compared with recipients with an eGFR ≥60 ml/min/1.73 $m^2$ and normal protein excretion (62) Although the integration of proteinuria and eGFR assessment has been shown to be a very good predictor of graft outcome (61, 63), more data must be collected before this combination can be advocated as a study endpoint in kidney transplantation clinical trials.

Finally, the causes of long-term graft failure are complex, and there are good arguments to capture this heterogeneity in more integrated composite scoring systems. For this, we refer to the paper in this supplement on surrogate endpoints (64).

## CONCLUSIONS

- Chronic renal graft dysfunction resembles CKD of native kidneys in many aspects:
  - Loss of nephrons, glomerulosclerosis, interstitial fibrosis, and tubular atrophy are pathological hallmarks of both.
  - Dysfunction is reflected as GFR loss, with or without proteinuria, ultimately leading to ESRD; however, ESRD is typically a late event and its use as an endpoint in clinical trials requires very large numbers of patients and prolonged follow-up.
- The EMA 2016 guideline recommended the time to a predefined and justified loss in GFR, such as 50%, as an endpoint in secondary prevention trials.
- Since the guideline was released, additional literature has been published on the choice of endpoints in CKD, and various endpoints have been used in clinical trials.
  - Many studies in CKD and kidney transplantation show that a change in eGFR (MDRD or CKD-EPI formulas) is strongly associated with ESRD.
  - It is increasingly advocated to use eGFR trajectories as endpoints in intervention trials in CKD. A caveat is the occurrence of an acute treatment effect that hampers use of the GFR trajectory; therefore, in kidney transplantation, special consideration should be given to studies including initiation or discontinuation of calcineurin inhibitors.
  - No studies convincingly demonstrate that measured GFR is a better predictor of ESRD than eGFR.
- In studies including patients with advanced-stage graft dysfunction (eGFR <45 ml/min/1.73 $m^2$) and/or rapid decline of GFR (>5 ml/min/1.73 $m^2$ per year), a composite endpoint consisting of a 30%–40% decline in eGFR or ESRD occurrence is both robust and feasible.
- In studies aimed at improving the lifespan of a transplanted kidney with more conserved renal function (eGFR >45 ml/min/1.73 $m^2$), eGFR time course (expressed as slope) should be accepted as surrogate endpoint, provided that the following limitations are considered:
  - Use of the chronic eGFR slope is inappropriate when a treatment has acute effects on GFR that are relatively large compared with expected chronic effects. In such cases, use of the total eGFR slope is generally favored.
  - Creatinine-based formulas to estimate GFR can be imprecise when there are non-GFR determinants of the creatinine concentration, such as use of drugs that inhibit tubular secretion [trimethoprim] or changes in muscle mass due to corticosteroid treatment.
  - Accuracy of cystatin C-based formulas to estimate eGFR can be influenced by corticosteroid use.
- While proteinuria/albuminuria appears to be a useful surrogate endpoint for CKD progression, especially in diabetes and IgA nephropathy, more research must be undertaken before proteinuria (or the combination of eGFR and proteinuria) can be advocated as an endpoint in studies in kidney transplant recipients.

## Scientific Advice From the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA) Regarding These Conclusions

This paper provides a contemporary discussion of graft functional parameters as primary endpoints in clinical trials as endpoints for clinical trials in kidney transplantation. ESOT has come to the following conclusions:

- The CHMP agreed that endpoints to assess efficacy of medicinal products to slow progression of chronic renal insufficiency (3) can be adopted to trials of kidney transplantation.
  - These include hard clinical endpoints (incidence of ESRD and renal/overall survival), proportional decrease in eGFR, and annual decrease in eGFR (slope).
- The CHMP agreed that conceptual approaches used to assess efficacy endpoints for dysfunction can be extrapolated to kidney transplantation, as far as the concomitant medications and diseases are comparable:
  - The impact of additional nephrotoxicity (e.g., in cases of CNI or viral nephropathy due to over-immunosuppression) should be delineated from lower potential to preserve functional efficacy.
- The CHMP stated that multiple definitions of efficacy endpoints using GFR have been proposed: the most conservative of these is the 57% reduction in GFR, reflecting doubling of serum creatinine; more recently, lesser degrees of proportional reduction in GFR have been proposed.
- The CHMP agreed that several publications advocate use of eGFR slope as a surrogate for clinical outcome in kidney disease trials, with the following notes:
  - eGFR slope should not replace any of the aforementioned GFR-based surrogate endpoints, but should rather be understood as an additional tool to estimate renal benefit; choice of GFR-based endpoint will depend on baseline rate of GFR decline, feasibility issues (e.g., disease prevalence, estimated efficacy of the medicinal product); GFR-based endpoints could also be used to address efficacy in trials of renal transplantation.

- Annualized loss of GFR does not meet all criteria for a valid surrogate endpoint, but (properly defined) is considered as a valuable measure of efficacy in addition to the currently accepted hard clinical endpoints (incidence of ESRD and renal/overall survival). Loss of GFR is most often assessed through serial estimates of GFR (eGFR) but can also be assessed as proportional reduction in GFR (30%–57%).
- The main purpose for a slope-based endpoint in the assessment of therapy in CKD is when feasibility is an issue using standard endpoints, as might be the case in studies of rare and/or early kidney disease. In addition, the value of GFR slope in assessing a medicinal product may be evident during early clinical stages, i.e., in exploratory studies.
- Several issues should be addressed before determining the acceptability of GFR slope as an efficacy assessment in phase III studies to support market access authorization. The main prerequisites are:
  - Low prevalence of the condition as reflected in the target population. It may not be feasible to determine efficacy using standard endpoints in rare diseases or in subpopulations of more common diseases.
  - Slow rate of progression of the kidney disease in the target population. Obviously, assessment of efficacy using standard endpoints may be feasible if rate of progression is rapid.
- Other important considerations:
  - Linearity of the slope. A final decision cannot be made without detailed understanding of nature and patterns of acute and chronic phases of the GFR slope in the context of the specific kidney disease and the pharmacological actions of the investigational compound [see (3, 48)].
  - Suitability of GFR slope as the primary endpoint should be determined on a case-by-case basis. This includes the assessment of how best to analyse efficacy based on eGFR slopes, especially in the context of issues around a possible acute drug effect and linearity assumptions of the GFR measurements.
  - Intercurrent events and confounding. As for any endpoint assessment, development should consider clear definitions of intercurrent events (e.g., death, concomitant medication, treatment discontinuation) and missing data, and a clear understanding of how to handle these issues on a case-by-case basis.
  - Clinically relevant magnitude of effect size. Clinical significance of the proposed difference in slope progressions between treatment arms (active or placebo) should be defined for the specific development. An annual difference of 1 ml/min/1.73 $m^2$ for 2 years has been proposed as a clinically significant difference compared with placebo (50). This difference was not accepted as a general cut-off by the CHMP and should be justified for the target population based on baseline GFR and rate of progression of the underlying disease and study population.
  - Efficacy should be supported by other clinical measures, e.g. a second study or other endpoints, most often the standard renal endpoints. The benefit as assessed by these endpoints should be in the same direction as that of the GFR slope.
- The CHMP agreed that proteinuria/albuminuria is of limited value as an endpoint in kidney transplantation.

## AUTHOR CONTRIBUTIONS

This article was developed from the Broad Scientific Advice request submitted to the EMA/CHMP by ESOT in 2020: interactions between EMA and ESOT regarding this request began in 2016. For the present article, through virtual and face-to-face discussions, the working group on functional endpoints in kidney transplantation developed the ESOT position on the core question "Does the CHMP agree with the proposed definitions of graft (dys)function in kidney transplantation, and the recommendations for parameters that could be used as primary endpoints in clinical trial settings". The Centre for Evidence in Transplantation (CET) supported specific data extraction requests: these literature searches formed the basis of evidence used in the advice request and the present article. Input into the working group's output was provided from all ESOT members involved in the advice request process. The present article was adapted by LH and MN from the final Broad Scientific Advice request submission (June 2020), presentation documents and minutes of the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020). The revised draft was reviewed, finalized, and approved by all co-authors before submission.

## FUNDING

This work was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

LH reports speaker fees from Astellas, consultancy and research support from Chiesi, consultancy for Novartis, and research support from Sandoz. KB has received honoraria and/or research funding from Alexion, Astellas, Bristol Myers Squibb, Chiesi, Fresenius, Hansa, Hexal, Merck, Novartis, Otsuka, Pfizer, Roche, Sandoz, Siemens, and Veloxis. LF has received honoraria and/or research funding from Astellas, Chiesi, Hansa, and Novartis. JG consults for Sanofi. UH has received grants/research support from Baxter, Chiesi, and Neovii; speakers' bureaux/honoraria from Chiesi and Hansa; and consulting fees from Astellas, Hansa, Neovii, Novartis, and Teva. DH has received lecture fees and consulting fees from Astellas, Chiesi, MedinCell, Novartis, and Vifor; and grant support (paid to institution) from Astellas, Bristol Myers Squibb, and Chiesi. RO has received grants/research support from Amgen,

Astellas, and Chiesi; and speakers' bureaux/honoraria from Amgen, Astellas, Chiesi, Hansa, Neovii, Novartis, and Teva. MR has received lecture fees from Astellas and Chiesi; and research grant support (paid to institution) from Astellas and Chiesi for investigator initiated studies. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. Naesens M, Kuypers DRJ, De Vusser K, Evenepoel P, Claes K, Bammens B, et al. The Histology of Kidney Transplant Failure. *Transplantation* (2014) 98: 427–35. doi:10.1097/tp.0000000000000183
2. McCaughan JA, Courtney AE, Maxwell AP. Estimated Glomerular Filtration Rate Decline as a Predictor of Dialysis in Kidney Transplant Recipients. *Am J Nephrol* (2014) 39:297–305. doi:10.1159/000360426
3. European Medicines Agenc. Guideline on the Clinical Investigation of Medicinal Products to Prevent Development/slow Progression of Chronic Renal Insufficiency (2016). Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-prevent-development/slow-progression-chronic-renal-insufficiency_en.pdf (Accessed December 13, 2021).
4. Levey A, Inker L. Assessment of Glomerular Filtration Rate in Health and Disease: a State of the Art Review. *Clin Pharmacol Ther* (2017) 102:405–19. doi:10.1002/cpt.729
5. Risch L, Herklotz R, Blumberg A, Huber AR. Effects of Glucocorticoid Immunosuppression on Serum Cystatin C Concentrations in Renal Transplant Patients. *Clin Chem* (2001) 47:2055–9. doi:10.1093/clinchem/47.11.2055
6. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D. A More Accurate Method to Estimate Glomerular Filtration Rate from Serum Creatinine: A New Prediction Equation. *Ann Intern Med* (1999) 130:461. doi:10.7326/0003-4819-130-6-199903160-00002
7. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, Feldman HI, et al. A New Equation to Estimate Glomerular Filtration Rate. *Ann Intern Med* (2009) 150:604–12. doi:10.7326/0003-4819-150-9-200905050-00006
8. Santos J, Martins LS. Estimating Glomerular Filtration Rate in Kidney Transplantation: Still Searching for the Best Marker. *World J Nephrol* (2015) 4:345. doi:10.5527/wjn.v4.i3.345
9. Salvador CL, Hartmann A, Åsberg A, Bergan S, Rowe AD, Mørkrid L. Estimating Glomerular Filtration Rate in Kidney Transplant Recipients: Comparing a Novel Equation with Commonly Used Equations in This Population. *Transplant Direct* (2017) 3:e332. doi:10.1097/txd.0000000000000742
10. Porrini E, Ruggenenti P, Luis-Lima S, Carrara F, Jiménez A, de Vries APJ, et al. Estimated GFR: Time for a Critical Appraisal. *Nat Rev Nephrol* (2019) 15: 177–90. doi:10.1038/s41581-018-0080-9
11. Vincenti F, Charpentier B, Vanrenterghem Y, Rostaing L, Bresnahan B, Darji P, et al. A Phase III Study of Belatacept-Based Immunosuppression Regimens versus Cyclosporine in Renal Transplant Recipients (BENEFIT Study). *Am J Transplant* (2010) 10:535–46. doi:10.1111/j.1600-6143.2009.03005.x
12. Durrbach A, Pestana JM, Pearson T, Vincenti F, Garcia VD, Campistol J, et al. A Phase III Study of Belatacept versus Cyclosporine in Kidney Transplants from Extended Criteria Donors (BENEFIT-EXT Study). *Am J Transplant* (2010) 10:547–57. doi:10.1111/j.1600-6143.2010.03016.x
13. Weir MR, Mulgaonkar S, Chan L, Shidban H, Waid TH, Preston D, et al. Mycophenolate Mofetil-Based Immunosuppression with Sirolimus in Renal Transplantation: a Randomized, Controlled Spare-The-Nephron Trial. *Kidney Int* (2011) 79:897–907. doi:10.1038/ki.2010.492
14. Mjörnstedt L, Schwartz Sørensen S, von zur Mühlen B, Jespersen B, Hansen JM, Bistrup C, et al. Renal Function Three Years after Early Conversion from a Calcineurin Inhibitor to Everolimus: Results from a Randomized Trial in Kidney Transplantation. *Transpl Int* (2015) 28:42–51. doi:10.1111/tri.12437
15. Rummo OO, Carmellini M, Rostaing L, Oberbauer R, Christiaans MHL, Mousson C, et al. ADHERE: Randomized Controlled Trial Comparing Renal Function Inde Novokidney Transplant Recipients Receiving Prolonged-Release Tacrolimus Plus Mycophenolate Mofetil or Sirolimus. *Transpl Int* (2017) 30:83–95. doi:10.1111/tri.12878
16. Ku E, Xie D, Shlipak M, Hyre Anderson A, Chen J, Go AS, et al. Change in Measured GFR versus eGFR and CKD Outcomes. *J Am Soc Nephrol* (2016) 27: 2196–204. doi:10.1681/asn.2015040341
17. Perico N, Ruggenenti P, Perna A, Caroli A, Trillini M, Sironi S, et al. Octreotide-LAR in Later-Stage Autosomal Dominant Polycystic Kidney Disease (ALADIN 2): a Randomized, Double-Blind, Placebo-Controlled, Multicenter Trial. *PLoS Med* (2019) 16:e1002777. doi:10.1371/journal.pmed.1002777
18. Lambers Heerspink HJ, Gansevoort RT, Brenner BM, Cooper ME, Parving HH, Shahinfar S, et al. Comparison of Different Measures of Urinary Protein Excretion for Prediction of Renal Events. *J Am Soc Nephrol* (2010) 21:1355–60. doi:10.1681/asn.2010010063
19. Ying T, Clayton P, Naresh C, Chadban S. Predictive Value of Spot versus 24-hour Measures of Proteinuria for Death, End-Stage Kidney Disease or Chronic Kidney Disease Progression. *BMC Nephrol* (2018) 19:55. doi:10.1186/s12882-018-0853-1
20. Zhao Y-f., Zhu L, Liu L-j., Shi S-f., Lv J-c., Zhang H. Measures of Urinary Protein and Albumin in the Prediction of Progression of IgA Nephropathy. *Clin J Am Soc Nephrol* (2016) 11:947–55. doi:10.2215/cjn.10150915
21. Talreja H, Akbari A, White CA, Ramsay TO, Hiremath S, Knoll G. Predicting Kidney Transplantation Outcomes Using Proteinuria Ascertained from Spot Urine Samples versus Timed Urine Collections. *Am J Kidney Dis* (2014) 64: 962–8. doi:10.1053/j.ajkd.2014.07.027
22. Knight SR, Hussain S. Variability in the Reporting of Renal Function Endpoints in Immunosuppression Trials in Renal Transplantation: Time for Consensus. *Clin Transplant* (2016) 30:1584–90. doi:10.1111/ctr.12861
23. Budde K, Rath T, Sommerer C, Haller H, Reinke P, Witzke O, et al. Renal, Efficacy and Safety Outcomes Following Late Conversion of Kidney Transplant Patients from Calcineurin Inhibitor Therapy to Everolimus: the Randomized APOLLO Study. *Clin Nephrol* (2015) 83:11–21. doi:10.5414/cn108444
24. Gatault P, Bertrand D, Büchler M, Colosio C, Hurault de Ligny B, Weestel P-F, et al. Eight-year Results of the Spiesser Study, a Randomized Trial Comparingde Novosirolimus and Cyclosporine in Renal Transplantation. *Transpl Int* (2016) 29:41–50. doi:10.1111/tri.12656
25. Tedesco-Silva H, Peddi VR, Sánchez-Fructuoso A, Marder BA, Russ GR, Diekmann F, et al. Open-label, Randomized Study of Transition from Tacrolimus to Sirolimus Immunosuppression in Renal Allograft Recipients. *Transpl Direct* (2016) 2:e69. doi:10.1097/txd.0000000000000579
26. Knoll GA, Fergusson D, Chassé M, Hebert P, Wells G, Tibbles LA, et al. Ramipril versus Placebo in Kidney Transplant Patients with Proteinuria: a

Multicentre, Double-Blind, Randomised Controlled Trial. *Lancet Diabetes Endocrinol* (2016) 4:318–26. doi:10.1016/s2213-8587(15)00368-x
27. De Fijter JW, Holdaas H, Øyen O, Sanders J-S, Sundar S, Bemelman FJ, et al. Early Conversion from Calcineurin Inhibitor- to Everolimus-Based Therapy Following Kidney Transplantation: Results of the Randomized ELEVATE Trial. *Am J Transplant* (2017) 17:1853–67. doi:10.1111/ajt.14186
28. 3C Study Collaborative Group. Campath, Calcineurin Inhibitor Reduction, and Chronic Allograft Nephropathy (The 3C Study) - Results of a Randomized Controlled Clinical Trial. *Am J Transplant* (2018) 18:1424–34. doi:10.1111/ajt.14619
29. Eskandary F, Regele H, Baumann L, Bond G, Kozakowski N, Wahrmann M, et al. A Randomized Trial of Bortezomib in Late Antibody-Mediated Kidney Transplant Rejection. *J Am Soc Nephrol* (2018) 29:591–605. doi:10.1681/asn.2017070818
30. Pascual J, Berger SP, Witzke O, Tedesco H, Mulgaonkar S, Qazi Y, et al. Everolimus with Reduced Calcineurin Inhibitor Exposure in Renal Transplantation. *J Am Soc Nephrol* (2018) 29:1979–91. doi:10.1681/asn.2018010009
31. Neal B, Perkovic V, Matthews DR, Mahaffey KW, Fulcher G, Meininger G, et al. Rationale, Design and Baseline Characteristics of the CANagliflozin cardioVascular Assessment Study-Renal (CANVAS-R): a Randomized, Placebo-Controlled Trial. *Diabetes Obes Metab* (2017) 19:387–93. doi:10.1111/dom.12829
32. Wanner C, Inzucchi SE, Lachin JM, Fitchett D, von Eynatten M, Mattheus M, et al. Empagliflozin and Progression of Kidney Disease in Type 2 Diabetes. *N Engl J Med* (2016) 375:323–34. doi:10.1056/nejmoa1515920
33. Coresh J, Turin TC, Matsushita K, Sang Y, Ballew SH, Appel LJ, et al. Decline in Estimated Glomerular Filtration Rate and Subsequent Risk of End-Stage Renal Disease and Mortality. *JAMA* (2014) 311:2518. doi:10.1001/jama.2014.6634
34. Levey AS, Inker LA, Matsushita K, Greene T, Willis K, Lewis E, et al. GFR Decline as an End point for Clinical Trials in CKD: a Scientific Workshop Sponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am J Kidney Dis* (2014) 64:821–35. doi:10.1053/j.ajkd.2014.07.030
35. Clayton PA, Lim WH, Wong G, Chadban SJ. Relationship between eGFR Decline and Hard Outcomes after Kidney Transplants. *J Am Soc Nephrol* (2016) 27:3440–6. doi:10.1681/asn.2015050524
36. Faddoul G, Nadkarni GN, Bridges ND, Goebel J, Hricik DE, Formica R, et al. Analysis of Biomarkers within the Initial 2 Years Posttransplant and 5-Year Kidney Transplant Outcomes. *Transplantation* (2018) 102:673–80. doi:10.1097/tp.0000000000002026
37. Bodine SC, Furlow JD. Glucocorticoids and Skeletal Muscle. *Adv Exp Med Biol* (2015) 872:145–76. doi:10.1007/978-1-4939-2895-8_7
38. Naesens M, Kuypers DRJ, Sarwal M. Calcineurin Inhibitor Nephrotoxicity. *Clin J Am Soc Nephrol* (2009) 4:481–508. doi:10.2215/cjn.04800908
39. Sommerer C, Duerr M, Witzke O, Lehner F, Arns W, Kliem V, et al. Five-year Outcomes in Kidney Transplant Patients Randomized to Everolimus with Cyclosporine Withdrawal or Low-Exposure Cyclosporine versus Standard Therapy. *Am J Transplant* (2018) 18:2965–76. doi:10.1111/ajt.14897
40. Smith-Palmer J, Kalsekar A, Valentine W. Influence of Renal Function on Long-Term Graft Survival and Patient Survival in Renal Transplant Recipients. *Curr Med Res Opin* (2014) 30:235–42. doi:10.1185/03007995.2013.855189
41. Kaplan B, Schold and J, Meier-Kriesche H-U. Poor Predictive Value of Serum Creatinine for Renal Allograft Loss. *Am J Transplant* (2003) 3:1560–5. doi:10.1046/j.1600-6135.2003.00275.x
42. Perkovic V, Jardine MJ, Neal B, Bompoint S, Heerspink HJL, Charytan DM, et al. Canagliflozin and Renal Outcomes in Type 2 Diabetes and Nephropathy. *N Engl J Med* (2019) 380:2295–306. doi:10.1056/nejmoa1811744
43. Torres IB, Reisaeter AV, Moreso F, Åsberg A, Vidal M, Garcia-Carro C, et al. Tacrolimus and Mycophenolate Regimen and Subclinical Tubulo-Interstitial Inflammation in Low Immunological Risk Renal Transplants. *Transpl Int* (2017) 30:1119–31. doi:10.1111/tri.13002
44. Cheung AK, Rahman M, Reboussin DM, Craven TE, Greene T, Kimmel PL, et al. Effects of Intensive BP Control in CKD. *J Am Soc Nephrol* (2017) 28:2812–23. doi:10.1681/asn.2017020148
45. Lambers Heerspink HJ, Weldegiorgis M, Inker LA, Gansevoort R, Parving H-H, Dwyer JP, et al. Estimated GFR Decline as a Surrogate End Point for Kidney Failure: A Post Hoc Analysis from the Reduction of End Points in Non-insulin-dependent Diabetes with the Angiotensin II Antagonist Losartan (RENAAL) Study and Irbesartan Diabetic Nephropathy Trial (IDNT). *Am J Kidney Dis* (2014) 63:244–50. doi:10.1053/j.ajkd.2013.09.016
46. Baigent C, Herrington WG, Coresh J, Landray MJ, Levin A, Perkovic V, et al. Challenges in Conducting Clinical Trials in Nephrology: Conclusions from a Kidney Disease-Improving Global Outcomes (KDIGO) Controversies Conference. *Kidney Int* (2017) 92:297–305. doi:10.1016/j.kint.2017.04.019
47. Levey AS, Gansevoort RT, Coresh J, Inker LA, Heerspink HL, Grams ME, et al. Change in Albuminuria and GFR as End Points for Clinical Trials in Early Stages of CKD: A Scientific Workshop Sponsored by the National Kidney Foundation in Collaboration with the US Food and Drug Administration and European Medicines Agency. *Am J Kidney Dis* (2020) 75:84–104. doi:10.1053/j.ajkd.2019.06.009
48. Holtkamp F, Gudmundsdottir H, Maciulaitis R, Benda N, Thomson A, Vetter T. Change in Albuminuria and Estimated GFR as End Points for Clinical Trials in Early Stages of CKD: a Perspective from European Regulators. *Am J Kidney Dis* (2020) 75:6–8. doi:10.1053/j.ajkd.2019.07.019
49. Grams ME, Sang Y, Ballew SH, Matsushita K, Astor BC, Carrero JJ, et al. Evaluating Glomerular Filtration Rate Slope as a Surrogate End point for ESKD in Clinical Trials: an Individual Participant Meta-Analysis of Observational Data. *J Am Soc Nephrol* (2019) 30:1746–55. doi:10.1681/asn.2019010008
50. Inker LA, Heerspink HJL, Tighiouart H, Levey AS, Coresh J, Gansevoort RT, et al. GFR Slope as a Surrogate End point for Kidney Disease Progression in Clinical Trials: a Meta-Analysis of Treatment Effects of Randomized Controlled Trials. *J Am Soc Nephrol* (2019) 30:1735–45. doi:10.1681/asn.2019010007
51. Greene T, Ying J, Vonesh EF, Tighiouart H, Levey AS, Coresh J, et al. Performance of GFR Slope as a Surrogate End point for Kidney Disease Progression in Clinical Trials: a Statistical Simulation. *J Am Soc Nephrol* (2019) 30:1756–69. doi:10.1681/asn.2019010009
52. Irish W, Nickerson P, Astor BC, Chong E, Wiebe C, Moreso F, et al. Change in Estimated GFR and Risk of Allograft Failure in Patients Diagnosed with Late Active Antibody-Mediated Rejection Following Kidney Transplantation. *Transplantation* (2021) 105:648–59. doi:10.1097/tp.0000000000003274
53. Cravedi P, Remuzzi G. Pathophysiology of Proteinuria and its Value as an Outcome Measure in Chronic Kidney Disease. *Br J Clin Pharmacol* (2013) 76:516–23. doi:10.1111/bcp.12104
54. Carrero JJ, Grams ME, Sang Y, Ärnlöv J, Gasparini A, Matsushita K, et al. Albuminuria Changes Are Associated with Subsequent Risk of End-Stage Renal Disease and Mortality. *Kidney Int* (2017) 91:244–51. doi:10.1016/j.kint.2016.09.037
55. Lambers Heerspink HJ, Kröpelin TF, Hoekman J, de Zeeuw D. Drug-induced Reduction in Albuminuria Is Associated with Subsequent Renoprotection: a Meta-Analysis. *J Am Soc Nephrol* (2015) 26:2055–64. doi:10.1681/asn.2014070688
56. Thompson A, Carroll K, A. Inker L, Floege J, Perkovic V, Boyer-Suavet S, et al. Proteinuria Reduction as a Surrogate End point in Trials of IgA Nephropathy. *Clin J Am Soc Nephrol* (2019) 14:469–81. doi:10.2215/cjn.08600718
57. Heerspink HJL, Greene T, Tighiouart H, Gansevoort RT, Coresh J, Simon AL, et al. Change in Albuminuria as a Surrogate Endpoint for Progression of Kidney Disease: a Meta-Analysis of Treatment Effects in Randomised Clinical Trials. *Lancet Diabetes Endocrinol* (2019) 7:128–39. doi:10.1016/S2213-8587(18)30314-0
58. Roodnat JI, Mulder PGH, Rischen-Vos J, van Riemsdijk IC, van Gelder T, Zietse R, et al. Proteinuria after Renal Transplantation Affects Not Only Graft Survival but Also Patient Survival. *Transplantation* (2001) 72:438–44. doi:10.1097/00007890-200108150-00014
59. Halimi J-M, Buchler M, Al-Najjar A, Laouad I, Chatelet V, Marlière J-F, et al. Urinary Albumin Excretion and the Risk of Graft Loss and Death in Proteinuric and Non-proteinuric Renal Transplant Recipients. *Am J Transplant* (2007) 7:618–25. doi:10.1111/j.1600-6143.2007.01665.x

60. Naesens M, Lerut E, Emonds M-P, Herelixka A, Evenepoel P, Claes K, et al. Proteinuria as a Noninvasive Marker for Renal Allograft Histology and Failure: an Observational Cohort Study. *J Am Soc Nephrol* (2016) 27:281–92. doi:10.1681/asn.2015010062
61. Weiner DE, Park M, Tighiouart H, Joseph AA, Carpenter MA, Goyal N, et al. Albuminuria and Allograft Failure, Cardiovascular Disease Events, and All-Cause Death in Stable Kidney Transplant Recipients: a Cohort Analysis of the FAVORIT Trial. *Am J Kidney Dis* (2019) 73:51–61. doi:10.1053/j.ajkd.2018.05.015
62. Ohkuma T, Jun M, Chalmers J, Cooper ME, Hamet P, Harrap S, et al. Combination of Changes in Estimated GFR and Albuminuria and the Risk of Major Clinical Outcomes. *Clin J Am Soc Nephrol* (2019) 14:862–72. doi:10.2215/cjn.13391118
63. Lam NN, Tonelli M, Lentine KL, Hemmelgarn B, Ye F, Wen K, et al. Albuminuria and Posttransplant Chronic Kidney Disease Stage Predict Transplant Outcomes. *Kidney Int* (2017) 92:470–8. doi:10.1016/j.kint.2017.01.028
64. Naesens M, Budde K, Hillbrand L, Oberbauer R, Bellini MI, Glotz D, et al. Surrogate Endpoints for Late Kidney Transplantation Failure. *Transpl Int* (2022). 35. doi:10.3389/ti.2022.10136

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10138

# Alloimmune Risk Stratification for Kidney Transplant Rejection

*Oriol Bestard[1†], Olivier Thaunat[2†], Maria Irene Bellini[3], Georg A. Böhmig[4], Klemens Budde[5], Frans Claas[6], Lionel Couzi[7], Lucrezia Furian[8], Uwe Heemann[9], Nizam Mamode[10], Rainer Oberbauer[4], Liset Pengel[11], Stefan Schneeberger[12] and Maarten Naesens[13]**

[1]Department of Nephrology and Kidney Transplantation, Vall d'Hebrón University Hospital, Barcelona, Spain, [2]Department of Transplantation, Nephrology, and Clinical Immunology, Edouard Herriot Hospital, Hospices Civils de Lyon, Lyon, France, [3]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy, [4]Department of Nephrology and Dialysis, Medical University of Vienna, Vienna, Austria, [5]Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany, [6]Eurotransplant Reference Laboratory, Department of Immunology, Leiden University Medical Center, Leiden, Netherlands, [7]Department of Nephrology, Transplantation and Dialysis, Bordeaux University Hospital, Bordeaux, France, [8]Kidney and Pancreas Transplantation Unit, University of Padua, Padua, Italy, [9]Department of Nephrology, Technical University of Munich, Munich, Germany, [10]Department of Transplantation, Guy's and St Thomas' NHS Foundation Trust, London, United Kingdom, [11]Centre for Evidence in Transplantation, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom, [12]Department of General, Transplant, and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria, [13]Department of Microbiology, Immunology, and Transplantation, KU Leuven, Leuven, Belgium

Different types of kidney transplantations are performed worldwide, including biologically diverse donor/recipient combinations, which entail distinct patient/graft outcomes. Thus, proper immunological and non-immunological risk stratification should be considered, especially for patients included in interventional randomized clinical trials. This paper was prepared by a working group within the European Society for Organ Transplantation, which submitted a Broad Scientific Advice request to the European Medicines Agency (EMA) relating to clinical trial endpoints in kidney transplantation. After collaborative interactions, the EMA sent its final response in December 2020, highlighting the following: 1) transplantations performed between human leukocyte antigen (HLA)-identical donors and recipients carry significantly lower immunological risk than those from HLA-mismatched donors; 2) for the same allogeneic molecular HLA mismatch load, kidney grafts from living donors carry significantly lower immunological risk because they are better preserved and therefore less immunogenic than grafts from deceased donors; 3) single-antigen bead testing is the gold standard to establish the repertoire of serological sensitization and is used to define the presence of a recipient's circulating donor-specific antibodies (HLA-DSA); 4) molecular HLA mismatch analysis should help to further improve organ allocation compatibility and stratify immunological risk for primary alloimmune activation, but without consensus regarding which algorithm and cut-off to use it is difficult to integrate information into clinical practice/study design; 5) further clinical validation of other immune assays, such as those measuring anti-donor cellular memory (T/B cell ELISpot assays) and non–HLA-DSA, is needed; 6) routine clinical tests that reliably measure innate immune alloreactivity are lacking.

**Keywords: alloimmune risk, crossmatch, high-risk transplantation, individualized immunosuppression, molecular HLA mismatch**

OPEN ACCESS

***Correspondence:**
Maarten Naesens
maarten.naesens@kuleuven.be

[†]These authors have contributed equally to this work

***Received:*** *21 October 2021*
***Accepted:*** *26 January 2022*
***Published:*** *20 May 2022*

***Citation:***
*Bestard O, Thaunat O, Bellini MI, Böhmig GA, Budde K, Claas F, Couzi L, Furian L, Heemann U, Mamode N, Oberbauer R, Pengel L, Schneeberger S and Naesens M (2022) Alloimmune Risk Stratification for Kidney Transplant Rejection. Transpl Int 35:10138. doi: 10.3389/ti.2022.10138*

## INTRODUCTION

Over time, donor and recipient profiles have changed substantially (1) modifying the risk of allograft rejection. Thus defining distinct alloimmune and non-alloimmune factors driving allograft rejection is greatly needed. For example, the proportion of sensitized [i.e., with circulating anti-human leukocyte antigen (HLA) antibodies] patients on kidney transplant waiting lists has gradually increased worldwide, because of both the implementation of highly sensitive immune assays to identify them and the increased proportion of retransplantations. In parallel, the number of expanded-criteria donors (ECD) or donors after circulatory death (DCD)—both groups that are often dominated by elderly people—may now exceed 50% in many transplant programmes. In such scenarios, it can be difficult to attain the excellent kidney transplantation outcomes observed for low-risk recipients and standard-criteria donors. However, low-risk donor and recipient is the usual pairing included in randomized controlled trials investigating new molecules and immunosuppression strategies. Given their real-world complexities, it would be useful to establish endpoints to identify clinically relevant and affordable improvements in outcome for distinct high-risk transplantation scenarios. This article presents evidence-based key determinants in immunological and non-immunological risk stratification, including but also extending beyond clinical research settings.

Technologies to assess alloimmune risk in transplant recipients have been developed and implemented in clinical practice, but further improvements to alloimmune risk stratification in kidney transplantation are needed. Such improvements would help to identify different subgroups of transplantation patients with distinct immune risks, which in turn would inform the development of clinical studies. Risk stratification is an essential first step toward personalized immunosuppression strategies for kidney transplant recipients.

Long-term immunosuppressive therapy may cause transplant recipients to experience various clinical events including cardiovascular disease, oncologic or metabolic complications, or opportunistic infections. Currently it is hard to individualize immunosuppressive therapy regimens to minimize the risk of such complications since data on individualization strategies remain limited and do not yet enable specific high-risk profiles to be identified (2). By contrast, the risk of allograft rejection (i.e., the immune-mediated destruction of transplanted organs), a major cause of graft loss, has been extensively investigated in large, retrospective population-based cohorts (3,4).

## STRATEGIES TO EVALUATE ALLOIMMUNE RISK

Alloantigens are unavoidably recognized by the kidney transplant recipient's adaptive immune system. However, the innate immune system—which is triggered by damage-associated molecular patterns (DAMP) released in the circulation, because of ischemia-reperfusion injury (IRI) immediately after transplantation—is necessary to prime the adaptive alloimmune response. DAMP are strong stimulators of the immune system (5,6). Immunological dogma holds that rejection requires effectors of the adaptive immune system, namely alloreactive cytotoxic T cells and donor-reactive B cells, which produce destructive donor-specific antibodies (DSA). Notably, a key feature of the adaptive immune system compared with the innate immune system is that the former generates antigen-specific memory effectors (i.e., memory T and B cells), which respond rapidly when the same antigen is re-encountered. Importantly, although this vision of rejection as being largely dependent on the ability of the adaptive immune system to discriminate between alloantigens (i.e., a process named allorecognition) largely remains dominant, independent reports from basic-research and early clinical studies suggest that some innate effectors (including monocytes and natural killer cells) are also capable of allorecognition, leading to previously overlooked types of "innate" rejection episodes (7–9) and interfering with the adaptive immune mechanisms at stakes in "classical" rejection episodes (10).

Two main strategies are used worldwide for immune-risk stratification before kidney transplantation (11). First, evaluation of HLA disparity between recipient and donor, which quantifies the risk that a "naïve" transplant candidate will develop a *de novo* alloimmune response over time, by recognizing foreign alloantigens. Secondly, identification of preformed circulating IgG antibodies against HLA in the recipient's serum, capable of lysing donor lymphocytes in a complement-dependent manner ("serological memory"); these antibodies are identified using a complement-dependent cell (CDC)-crossmatch assay. The latter approach aims to identify sensitized transplant candidates with *preformed* humoral alloimmunity, able to trigger complement cascade activation against the graft (i.e., preformed DSA responsible for rapid severe AMR and graft loss).

Advances in the characterization of donor/recipient HLA disparities at the molecular level, use of the flow crossmatch (FCXM) (in some centers), and novel and highly sensitive immunological tools to detect circulating IgG anti-HLA antibodies (whether complement binding or not), have substantially changed the landscape of immune-risk profiling.

## IMMUNOGENICITY OF KIDNEY GRAFT DEPENDS ON DONOR CHARACTERISTICS

Kidney transplants from donors who are elderly, ECD, ECD/DCD, or kidneys with pre-existing lesions have poorer prognoses than transplants from standard-criteria donors. Recipients of organs from high-risk donors tend to have poor renal function, with reduced medium-term graft survival (12,13). Such patients are also at high risk of delayed graft function (DGF) (14,15). These problems are not only related to the lower intrinsic quality of such organs but also to their highest level of immunogenicity (discussed more, below).

Immunosuppressive regimens administered to recipients of kidneys from ECD are adapted to avoid early acute rejection, which might worsen any pre-existing or ischemic injury of the graft; conversely, the goal of maintenance immunosuppression in such settings is to attenuate the long-term nephrotoxicity of calcineurin inhibitors. Indeed, reducing IRI in high-risk donors has been a major goal in kidney transplantation to minimize not only the risk of DGF but also to abrogate subsequent alloimmune activation favoring allograft rejection (16). Agents counteracting the effects of ischemia have been studied in selected populations, mainly by using donor/recipient risk indices to assess DGF risk (17). Interventional studies to prevent graft IRI have generally evaluated DGF occurrence as a qualitative phenomenon, although some also evaluated medium-term renal function. Given that IRI is a dynamic response to numerous molecular events, assessing DGF severity could help with the quantitative evaluation of the protective effects of anti-ischemic agents. Indeed, following discussions with the US Food and Drug Administration, this approach is in clinical investigation (ClinicialTrials.gov identifier: NCT02610296), with DGF severity (measured in terms of the number of dialysis sessions required in the first 30 days post transplantation, for participants starting dialysis on days 0–7) as the primary endpoint. However, DGF is unspecific and only partially relates to long-term graft function.

It is beyond the scope of the present paper to discuss ECD criteria and the definition of DGF as a potential endpoint in more detail, although it may also be a consequence of an early acute rejection episode; instead, we focus on the establishment of alloimmune risk in transplantation settings. Nevertheless, it should be noted that we consider DGF as a potentially meaningful endpoint for registration trials in transplantation.

## OPTIMIZING OUTCOMES FOR HIGHLY SENSITIZED PATIENTS

Compared with other candidates, highly sensitized patients have reduced access to kidney transplantation and worse allograft outcomes, mainly due to their high risk for antibody-mediated rejection (AMR) (18).

The broadness of anti-HLA sensitization is evaluated using calculated panel-reactive antibody (cPRA) testing, which for each candidate estimate the percentage of donors against whom he is likely to show DSA, thus ultimately determining the proportion of unacceptable donors for a given transplant candidate. Candidates with very high cPRA values (>90%) have a reduced chance of finding a suitable kidney donor (19).

Several strategies enhance access to transplantation in highly sensitized candidates. The best option is the transplantation of a kidney from an HLA-compatible living donor (20,21) which, in the absence of an HLA-identical sibling volunteering for donation, may be achieved through large paired-donor exchange pools.

In the absence of a compatible donor, in the United States, desensitization protocols are commonly used (22). Standard-of-care desensitization regimens are based on a combination of off-label agent usage and techniques that aim to reduce antibody titers transiently, such as administration of intravenous immunoglobulin (IVIg), rituximab, and pre- and/or post-transplant apheresis with plasma exchange or immunoadsorption. However, such approaches are not widely followed in Europe, for reasons including evidence of inferior outcomes compared with HLA-compatible transplantation and a lack of robust data demonstrating the superiority of these high risk costly procedures (20,23,24).

In the absence of an HLA-compatible living donor, three strategies exist to increase access to transplantation for highly sensitized candidates on deceased-donor transplant waiting lists in Europe, which consider the degree of sensitization in algorithms for organ allocation. First, in the United States and some European countries including Spain and France, Kidney Allocation Systems prioritize candidates with very high cPRA values (percentages differ among countries but are ≥95%). This has increased the transplantation rate among highly sensitized candidates to levels similar to those for other candidates; however, in extremely sensitized patients (cPRA ≥99.9%), transplantation rates remain significantly lower than rates for less-sensitized patients (25). Secondly, the Eurotransplant International Foundation developed the Acceptable Mismatch (AM) Program for highly sensitized patients in the 1980s. Between 1989 and 2017, over 2,500 patients were listed on the Program, 57% of whom received a donor kidney (26). The 10-years graft survival rate among recipients listed on the Eurotransplant AM program was comparable to that for less sensitized recipients (26). The AM strategy is also used outside the Eurotransplant Program. For example, since 2005 France has operated a national AM policy (27). The EUROpe-wide Strategy to enhance Transplantation of highly sensitized patients based on Acceptable HLA Mismatches (EUROSTAM) project has developed and tested a tool to evaluate opportunities for sharing kidneys across different countries; the aim of this initiative is to increase HLA-compatible transplantation rates and thus, improve outcomes (28). There is also a third option: desensitization can be undertaken, sometimes in combination with specific allocation programs, to facilitate transplantation in sensitized recipients with preformed DSA (and/or positive crossmatch) (29,30).

Despite the creation of these programs to increase access to transplantation for highly sensitized candidates, a substantial number of people may not benefit (31), especially those with cPRA ≥98%, who often remain wait-listed for many years. These transplant candidates may need different strategies to increase their level of access to organs. In this regard, access to transplantation might be considered as a discrete endpoint among highly sensitized candidates enrolled in studies investigating whether new therapeutic approaches help to improve transplantation rates. Of note, fair evaluation of desensitization strategies based on access to transplantation requires that future studies enrol patients with homogeneous humoral immunological risk (discussed below). Clear distinction appears to be mandatory between candidates with positive lymphocytotoxicity test (LCT) cross match against their donor (who require therapeutic action pre-transplantation to reduce the

titer of preformed DSA, to prevent hyperacute rejection) and candidates with lower DSA levels (positive FCXM and/or positive solid phase assay) and negative LCT, who can be transplanted without prior desensitization and only require adaptation of immunosuppression. In this regard, the transplantation rate alone is not a sufficient endpoint; only successful (e.g., rejection-free, good renal function) transplantations in the medium- or long-term (typically >10 years) should be considered. Highly sensitized transplant recipients are at high risk of developing AMR; they also have poor renal function and low graft survival rates (20,32). Although the relevance and impact of T cell-mediated rejection in these patients is lower compared with other transplantation groups, AMR with donor-specific anti-HLA antibodies exerts a detrimental effect on long-term graft survival (33,34). Hence, AMR could be a very suitable primary endpoint and surrogate for graft outcome in highly sensitized compared with HLA-compatible kidney transplant recipients.

The 2017 Banff conference described active AMR—which has several clinicopathological subtypes—as being indicative of ongoing disease activity. Active AMR is characterized by microvascular inflammation with or without graft remodeling; it is discussed further in the article by Becker et al. in the present Special Issue (35), and in Banff consensus publications (36,37).

# CONSIDERATIONS TO IMPROVE STRATIFICATION OF ALLOIMMUNE RISK ASSOCIATED WITH KIDNEY TRANSPLANTATION

## Immunological Profiling of the Graft

As mentioned earlier, the immune system does not mount a response against a protein antigen without an adjuvant, which provides the molecular signals necessary to prime immune-effector cells. In transplantation, several epidemiological studies report that kidneys from older or marginal donors (i.e., those with heightened levels of tissue inflammation) are more immunogenic than kidneys from donors with less inflammation—especially when given to young recipients, whose immune system is more responsive to simulation. For instance, IRI can be increased by factors such as DCD and long cold-ischemia time, and can lead to DAMP release (36,37), thus instigating alloimmune responses. There is no validated clinical tool to evaluate the confounding effect of the type of transplantation. However, we believe that experimental data clearly support the notion that transplantations performed with living-donor kidneys carry significantly lower immunological risk compared with transplantations performed with kidneys from DCD with similar antigenic load (38).

## Immunological Compatibility Between Donors and Recipients

The risk of the recipient's immune system developing a response against the donor kidney (allograft immunogenicity) depends on the number of potential antigenic targets, and the level of stimulation of the recipient's immune system by adjuvant molecules.

Large studies show that long-term kidney graft survival decreases with the number of HLA-mismatch antigens between donor and recipient (39,40). HLA mismatches used to be defined based on serological determination of A, B, and DR molecules in donor and recipient. Immunogenetic advances have improved the accuracy of donor/recipient HLA typing and revealed that not all HLA mismatches have the same impact on outcome. The immunological importance of a given HLA mismatch depends on the number of epitopes that can be recognized by the recipient's immune system (B and/or T cells) (41,42).

Progress in bioinformatics has facilitated integration of all these data to calculate the "epitope load"—a parameter that correlates much better with risk of developing dnDSA than simply counting the number of HLA mismatches (43–46). Epitope load is likely to play a key role in better stratifying the primary immunological risk associated with a specific transplantation and are associated with specific geographical regions that may not be extrapolated to a global level.

Furthermore, beyond mere quantity, not all epitopes appear to have the same immunogenic relevance: although qualitative aspects of epitopes are not well documented, publications have described certain physicochemical characteristics of different epitopes (47,48).

However, without consensus regarding which algorithm (and cut-off) should be used, and with the ongoing need for more comprehensive high-resolution (HR) HLA typing, it might be difficult to integrate such information immediately in clinical practice. The principle of diminishing epitope load can be implemented irrespective of a selected algorithm: using HLAMatchmaker, amino acid mismatching and physicochemical mismatch load were shown to have the same impact on outcome (49). Accordingly, we consider that in presence of complete HR donor/recipient HLA typing, or low resolution within biologically related pairs, transplantations performed with HLA-identical donors (in particular if donor and recipient are closely related, e.g., siblings) carry significantly lower immunological risk than those performed with donors of other HLA statuses.

### Anti-HLA Antibodies

Screening for anti-HLA antibodies is the cornerstone of immune-risk profiling in kidney transplant recipients. A positive CDC-crossmatch with donor cells is considered a contraindication to transplantation (unless desensitization is initiated before transplantation, a situation that is not discussed in the present article). CDC-crossmatch can also be assessed with a panel of different cells to evaluate the diversity of the recipient's serological memory against HLA molecules. A CDC-PRA test figure of >80% was historically used to define hyperimmunized patients and implies a lower access to transplantation (as their CDC-crossmatch with donor cells is more likely to be positive). Percentage of PRA has been applied for immune-risk stratification in large clinical trials. However, since the CDC-crossmatch only detects DSA that activate complement [a

characteristic that depends on the titer and specific biological characteristics of IgG (50)], some recipients with negative CDC-crossmatch and/or CDC-PRA might still reveal preformed DSA that also have a deleterious impact on graft survival through antibody-dependent cell cytotoxicity (21,51–53).

While a negative CDC-crossmatch with donor cells will remain a mandatory condition to perform transplantation, novel, or more sensitive techniques—such as FCXM and single-antigen bead (SAB) assays to detect alloantibodies—have been implemented to improve the screening of recipients for preformed DSA. These assays can detect circulating anti-HLA antibodies that can be a mixture of antibodies that do and do not fix complement and may harm the graft through antibody-dependent cell cytotoxicity and/or direct modulation of graft endothelial-cell biology (54), thus significantly improving the capacity of detecting pathogenic circulating DSAs. FCXM with donor cells is more sensitive than CDC and yields fewer false-positive results than solid-phase assay (55–57). However, it requires collection of the donor's cells and use of a cytometer, which is not available in all immunogenetic laboratories. Other limitations of FCXM include poor standardization, thresholds, and interpretation of test systems. Conversely, SAB assays are widely available, more standardized, and have better reproducibility than FCXM. SAB assays consist of microparticles coated with purified HLA antigens; if antibodies are present, a semiquantitative readout is provided. There are two commercial platforms: One Lambda® (Thermo Fisher Scientific, Canoga Park, CA, United States) and Immucor® (Immucor, Norcross, GA, United States) with rather good correlation and reliability between both assays (58). With some caveats, we propose that SAB assays should be the gold standard to establish the repertoire of serologic memory and define the presence of a recipient's circulating DSA. The caveats are that the results of SAB assays are semiquantitative, and they have certain technical limitations and interlaboratory variability: for example, it is necessary to prevent the artifact of complement interference by pre-treating serum (e.g., with EDTA or heat inactivation). In the absence of strong consensus to define the mean fluorescence intensity (MFI) cut-off that would indicate clinically relevant HLA antibodies, we suggest transplant physicians and immunologists should define the most appropriate cut-off for local circumstances. Establishing plausibility of the potential DSA, considering previous immunizing events, is a key factor to determine antibody positivity in the individual. Notably, several studies have reported that the ability of DSA identified by SAB to bind *ex vivo* donor cells in FCXM is a good predictor of subsequent AMR lesions and graft loss (in 50% and 30% of recipients, respectively) (21,52,57,59–61). Together, these data suggest that optimal performance of FCXM in identifying pathogenic DSA depends on both higher specificity (elimination of false positivity due to denatured HLA molecules on SAB) and lower sensitivity (so that only DSA with high titers are detected).

Utilizing the results of these analyses, transplant candidates could be categorized according to their level of immune sensitization at the time of transplantation. In alignment with the approach proposed by the STAR [Sensitization in Transplantation: Assessment of Risk (21,51,52)] and ENGAGE [EuropeaN Guidelines for the mAnagement of Graft rEcipients (11)] Working Groups, ESOT recommends differentiation of anti-HLA antibody status by categorizing patients (**Figure 1**). Using this approach, patients with HLA-DSA at the time of transplantation (day 0; group 3, **Figure 1**) would have a higher likelihood of post-transplant AMR and less favorable allograft outcomes than patients naïve for alloantigen (group 1, **Figure 1**). The situation is less clear for group 2 (patients with previous exposure to donor HLA antigens during a transplant or pregnancy or a history of HLA-DSA, but who are negative at the time of transplantation). A retrospective single-center study suggested a detrimental impact on outcome for these people (62). We recommend considering patients in group 2 as being at intermediate risk, because of the likely presence of cellular memory (discussed below) (**Figure 1**). There is ongoing discussion about the predictive value of the MFI in SAB testing of HLA-DSA (29) and whether the MFI could be used as a potential surrogate to estimate the HLA-DSA titer and, consequently, further refine the individual's risk of developing AMR in group 3 patients. However, technical aspects of the semiquantitative values should be considered: cut-offs may differ between centers.

## Non-HLA Antibodies

Not all antibodies implicated in kidney transplant rejection are directed against the HLA system. Accumulating clinical and experimental data indicate a deleterious role of antibodies, i.e., antibodies directed against graft antigens other than allogeneic HLA molecules (63,64). Of note, the nature of these antibodies, i.e., whether they are auto and/or alloreactive, remains currently unclear. In this regard, the demonstration that genetic mismatch of non-HLA haplotypes coding for transmembrane or secreted proteins is associated with an increased risk of functional graft loss, independently of HLA incompatibility, suggests that non-HLA antibodies could be alloreactive (65). Furthermore, this literature suggests there is enormous diversity among potential antigenic targets, complicating the detection of non–HLA-DSA. Until consensus is established, and in the absence of a validated assay (and cut-off value), we do not recommend that non–HLA-DSA are considered when evaluating the immunological risk for a transplantation.

## Adaptive Cellular Memory

In addition to immunological assessment of antibodies, screening for adaptive cellular memory seems to be valuable. Although ELISpot assays can identify donor-reactive memory B and T cells in kidney transplant recipients, to date these assays have only been shown to predict transplant outcomes in small, underpowered, retrospective studies (66–68). Standardization and cross-validation of the donor-specific T cell ELISpot assay between laboratories has been performed (69,70) and translation to clinical settings was attempted in a multicenter, randomized interventional trial (71). According to this study, rates of T cell-mediated rejection were significantly higher in patients with preformed donor-reactive T cell frequencies compared with other patients. However, these cell-based

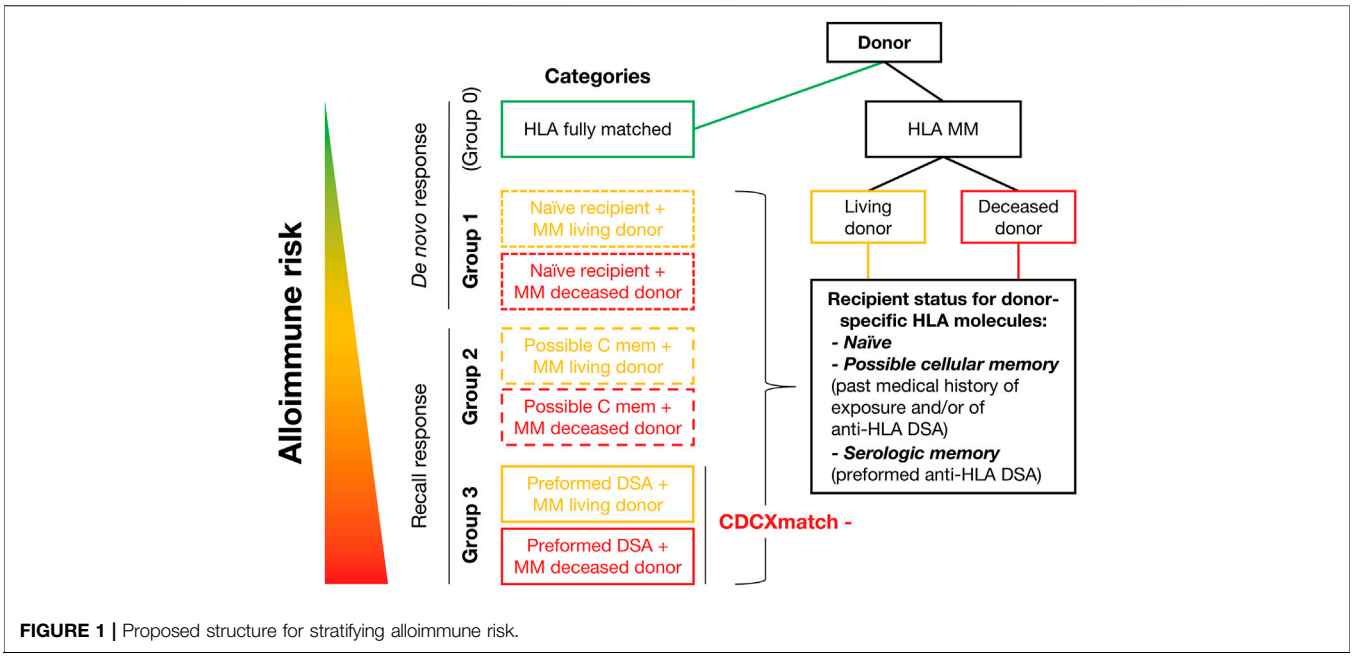


**FIGURE 1 |** Proposed structure for stratifying alloimmune risk.

assays need further evaluation of reproducibility before widespread clinical use.

Although we do not recommend implementation of T/B cell ELISpot assays in routine clinical practice, determining the presence of an adaptive cellular memory in the recipient seems important, to stratify the immunological risk of a specific transplantation according to immune sensitization status. This can be done by establishing the recipient's pregnancy history (to identify the father's HLA type), transplantation history (to identify the previous donor's HLA type), and transfusion history (red blood cells and platelets, although the profile of sensitization is usually complicated to assess) (72). While we acknowledge that this information might not be obtained in many cases and does not necessarily imply the presence of an effector anti-donor alloimmune response at the cellular level, in some specific transplant scenarios (e.g., living donor kidney transplantation) this information might be possible to retrieve and may help to better understand potential immunological events occurring during the early phases post-transplantation, underscoring a preformed recall anti-donor alloimmune response. A patient with a history of anti-HLA antibodies that are undetectable in the circulation should be considered likely to have memory B and/or T cells against these HLA antigens, especially those against previously recognized alloantigens. While there is no evidence on how to specifically manage these patients, such situations require special attention.

### Innate Immune Effectors

Finally, as well as participating in graft damage on recruitment by adaptive effectors, innate immune effectors might be able to recognize allogeneic non-self (7–9). While we wait for experimental studies to translate into robust clinical findings, and reliable assays are developed to guide decisions, we do not recommend that innate immune alloreactivity is evaluated in routine clinical practice.

## CONCLUSIONS

The following is our proposal for alloimmune risk stratification in CDC-negative kidney transplantation.

- Transplantation performed with an HLA-identical donor carries a significantly lower immunological risk than transplantation from a donor of another HLA status.
- For the same allogeneic eplet load, grafts from living donors, which are better preserved and are therefore less immunogenic than grafts from deceased donors, are associated with a lower immunological risk.
- Patients with anti-donor serological memory at the time of or short time before transplantation (i.e., those with the presence of HLA-DSA) should be clearly differentiated from the others:
  - Patients with donor reactivity are likely to have immune reactions to the allograft, with a heightened risk of post-transplant AMR and poor allograft outcome.
  - SAB testing is the gold standard to establish the repertoire of serologic memory and define the presence of a recipient's circulating anti-HLA DSA.
    - Local transplant physicians and immunologists should determine the appropriate cut-off point, with a focus on plausibility of immunization history.
    - In absence of clinical validation, non–HLA-DSA routine screening assays should not be considered when evaluating immunological risk of a transplantation.

- Using SAB testing, three risk groups can be identified (patients with non–donor-specific HLA antibodies and no previous exposure to donor antigen are considered as naïve patients):
  - Group 1: Patients with no signs of anti-HLA immune sensitization at any time point (very low risk).
  - Group 2: Patients with previous exposure to donor antigens or history of HLA-DSA positivity, but without HLA-DSA at time of transplantation (intermediate risk due to likely presence of memory T and/or B alloimmune response): T/B ELISpot assays could identify anti-donor memory cells, but without clinical validation these assays should not be considered when evaluating immunological risk of a transplantation.
  - Group 3: Patients with HLA-DSA at time of transplantation (high risk). There is ongoing discussion on the utility of MFI in SAB or FCXM as a potential surrogate to estimate the DSA titer and for individual risk stratification.
  - Molecular HLA mismatch analysis is likely to play a future role in better allocating more compatible allografts, as well as in stratifying the primary alloimmune risk. However, in the absence of consensus regarding what algorithm (and which cut-off) should be used to quantify the eplet load and whether the quality of eplet should also be considered, it is difficult to integrate such information immediately into clinical practice and clinical trial design. Further consensus building is necessary.
  - Although it is a fast-evolving field, no reliable test is currently available to measure innate immune alloreactivity in routine clinical practice.

## Scientific Advice From the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA) Regarding these Conclusions

- The CHMP agreed that several important issues need to be considered in assessing the alloimmune risk following kidney transplantation. These include general characteristics of the recipient and donor, as well as issues related to the transplanted organ and issues requiring further studies and/or consensus before adapting into general guidelines.
- Several of these factors and issues are already discussed in the EMA guideline CHMP/EWP/263148/06 (66). It is agreed that high-risk populations should be distinguished based on 1) greater risk of clinical events and 2) the need for different immunosuppression intensity.
  - Regarding the immunological risk related to the donated kidney, the CHMP agreed that the number of antigenic targets on the donated organ and "adjuvantation" affect the outcome of transplantation. Some of these issues will be addressed by the type of organ transplanted (ischemia time, HLA mismatch, living donation vs. ECD, DCD etc.), which, depending on the study design, can be used for stratification.
  - Regarding improving stratification of the recipients based on immunological profiling, the CHMP agreed that:
    - A positive CDC-crossmatch detects only DSA that activate complement. For risk stratification, this is not ideal, as DSA may still be present with deleterious impact on graft survival.
    - Other tests are more sensitive, such as the FCXM and SAB assays. ESOT proposes to use the SAB assay as the gold standard to define the presence of recipient's circulating DSA. The preference of SAB is advocated for sensitive anti-HLA DSA based on wide availability in practice. No data were submitted to support this conclusion. Furthermore, no definitive metrics are proposed (e.g., MFI cut-off values), leaving the cut-off definitions to local transplant physicians and immunologists. This flexibility of defining cut-offs in clinical practice is acknowledged. However, this raises issues for external validity of study results when the proposed metrics are not generally accepted.
    - Innate immune effects and cell-based assays addressing cellular memory need further evaluation before widespread clinical use and validation before application in clinical trials.
  - The CHMP stated that the classification in three risk categories based on the HLA antibody profiles is interesting and could be acceptable, if a general consensus in the transplant community supports the classification. Also, the definition of cut-offs to define anti-HLA positivity requires more work. Currently, for individual applications basis.
  - Finally, the CHMP stated that the stratification factors to be used in individual studies should reflect the goal and the size of the study.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication. This article is one of several developed from the Broad Scientific Advice request, submitted to the EMA/CHMP by ESOT in 2020. For the present article, through virtual and face-to-face discussions, the working group on risk stratification in kidney transplantation developed the ESOT position on the core question "Does CHMP agree with the proposed specific risk profiles for kidney transplantation which determine background risk of rejection associated with immunosuppressive therapy?" The Centre for Evidence in Transplantation provided support with specific data extraction requests: these literature searches formed the basis of evidence used in the advice request and the present article. Input into the working group's output was provided from all ESOT members involved in the advice request process. The present article was adapted by MN from

the final Broad Scientific Advice request submission (June 2020), presentation documents and minutes of the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020). The first draft of the article was reviewed by OB and OT; the revised draft was reviewed, finalized, and approved by all co-authors before submission for publication.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

OB has received research funding from Chiesi and served as adviser for Hansa Biopharma. OT has received research funding from bioMérieux, Bristol Myers Squibb, and Immucor; and has consultancy agreements with Biotest and Novartis. GB has received honoraria and/or research funding from Astellas, CareDx, CSL Behring, Fresenius, Hansa, Neovii, and Vitaeris. KB has received honoraria and/or research funding from Alexion, Astellas, Bristol Myers Squibb, Chiesi, Fresenius, Hansa, Hexal, Merck, Novartis, Otsuka, Pfizer, Roche, Sandoz, Siemens, and Veloxis. FC is a scientific adviser for GenDx and Immucor. LF has received honoraria and/or research funding from Astellas, Chiesi, Hansa, and Novartis. UH has received grants/research support from Baxter, Chiesi, and Neovii; speakers' bureaux/honoraria from Chiesi and Hansa; and consulting fees from Astellas, Hansa, Neovii, Novartis, and Teva. NM has received honoraria from Hansa, Chiesi, Novartis, and Takeda. RO has received grants/research support from Amgen, Astellas, and Chiesi; and speakers' bureaux/honoraria from Amgen, Astellas, Chiesi, Hansa, Neovii, Novartis, and Teva. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. Eurotransplant. *Statistics*. Available from: https://www.eurotransplant.org/statistics/statistics-library/ (Accessed January 15, 2021).
2. Hricik DE, Formica RN, Nickerson P, Rush D, Fairchild RL, Poggio ED, et al. Adverse Outcomes of Tacrolimus Withdrawal in Immune-Quiescent Kidney Transplant Recipients. *J Am Soc Nephrol* (2015) 26:3114–22. doi:10.1681/asn.2014121234
3. Süsal C, Wettstein D, Döhler B, MorathRuhenstroth CA, Ruhenstroth A, Scherer S, et al. Association of Kidney Graft Loss with De Novo Produced Donor-specific and Non-donor-specific HLA Antibodies Detected by Single Antigen Testing. *Transplantation* (2015) 99:1976–80. doi:10.1097/tp.0000000000000672
4. Sellarés J, De Freitas DG, Mengel M, Reeve J, Einecke G, Sis B, et al. Understanding the Causes of Kidney Transplant Failure: the Dominant Role of Antibody-Mediated Rejection and Nonadherence. *Am J Transpl* (2012) 12:388–99. doi:10.1111/j.1600-6143.2011.03840.x
5. Todd JL, Palmer SM. Danger Signals in Regulating the Immune Response to Solid Organ Transplantation. *J Clin Invest* (2017) 127:2464–72. doi:10.1172/jci90594
6. Fuquay R, Renner B, Kulik L, McCullough JW, Amura C, Strassheim D, et al. Renal Ischemia-Reperfusion Injury Amplifies the Humoral Immune Response. *J Am Soc Nephrol* (2013) 24:1063–72. doi:10.1681/asn.2012060560
7. Dai H, Friday AJ, Abou-Daya KI, Williams AL, Mortin-Toth S, Nicotra ML, et al. Donor SIRPα Polymorphism Modulates the Innate Immune Response to Allogeneic Grafts. *Sci Immunol* (2017) 2:eaam6202. doi:10.1126/sciimmunol.aam6202
8. Koenig A, Chen C-C, Marçais A, Barba T, Mathias V, Sicard A, et al. Missing Self Triggers NK Cell-Mediated Chronic Vascular Rejection of Solid Organ Transplants. *Nat Commun* (2019) 10:5350. doi:10.1038/s41467-019-13113-5
9. Callemeyn J, Lamarthée B, Koenig A, Koshy P, Thaunat O, Naesens M. Allorecognition and the Spectrum of Kidney Transplant Rejection. *Kidney Int* (2021) S0085-2538(21):01143–1. doi:10.1016/j.kint.2021.11.029
10. Callemeyn J, Senev A, Coemans M, Lerut E, Sprangers B, Kuypers D, et al. Missing Self-Induced Microvascular Rejection of Kidney Allografts: a Population-Based Study. *J Am Soc Nephrol* (2021) 32:2070–82. doi:10.1681/asn.2020111558
11. Bestard O, Couzi L, Crespo M, Kessaris N, Thaunat O. Stratifying the Humoral Risk of Candidates to a Solid Organ Transplantation: a Proposal of the ENGAGE Working Group. *Transpl Int* (2021) 34:1005–18. doi:10.1111/tri.13874
12. Querard A-H, Foucher Y, Combescure C, Dantan E, Larmet D, Lorent M, et al. Comparison of Survival Outcomes between Expanded Criteria Donor and Standard Criteria Donor Kidney Transplant Recipients: a Systematic Review and Meta-Analysis. *Transpl Int* (2016) 29:403–15. doi:10.1111/tri.12736
13. Montero N, Toapanta N, Pallarès N, Crespo M, Diekmann F, Guirado L, et al. Deciphering Transplant Outcomes of Expanded Kidney Allografts Donated after Controlled Circulatory Death in the Current Transplant Era. A Call for Caution. *Transpl Int* (2021) 34:2494–506. doi:10.1111/tri.14131
14. Cardinal H, Lamarche F, Grondin S, Marsolais P, Lagacé AM, Duca A, et al. Organ Donor Management and Delayed Graft Function in Kidney Transplant Recipients: A Multicenter Retrospective Cohort Study. *Am J Transpl* (2019) 19:277–84. doi:10.1111/ajt.15127
15. Demiselle J, Augusto J-F, Videcoq M, Legeard E, Dubé L, Templier F, et al. Transplantation of Kidneys from Uncontrolled Donation after Circulatory Determination of Death: Comparison with Brain Death Donors with or without Extended Criteria and Impact of Normothermic Regional Perfusion. *Transpl Int* (2016) 29:432–42. doi:10.1111/tri.12722
16. Summers DM, Johnson RJ, Allen J, Fuggle SV, Collett D, Watson CJ, et al. Analysis of Factors that Affect Outcome after Transplantation of Kidneys Donated after Cardiac Death in the UK: a Cohort Study. *The Lancet* (2010) 376:1303–11. doi:10.1016/s0140-6736(10)60827-6

17. Wlodek E, Kirkpatrick RB, Andrews S, Noble R, Schroyer R, Scott J, et al. A Pilot Study Evaluating GSK1070806 Inhibition of Interleukin-18 in Renal Transplant Delayed Graft Function. *PLoS One* (2021) 16:e0247972. doi:10.1371/journal.pone.0247972
18. Bradley JA, Baldwin WM, Bingaman A, Ellenrieder C, Gebel HM, Glotz D, et al. Antibody-Mediated Rejection-An Ounce of Prevention Is Worth a Pound of Cure. *Am J Transpl* (2011) 11:1131–9. doi:10.1111/j.1600-6143.2011.03581.x
19. Kransdorf EP, Pando MJ. Calculated Panel Reactive Antibody with Decimals: a Refined Metric of Access to Transplantation for Highly Sensitized Candidates. *Hum Immunol* (2017) 78:252–6. doi:10.1016/j.humimm.2016.12.009
20. Manook M, Koeser L, Ahmed Z, Robb M, Johnson R, Shaw O, et al. Post-listing Survival for Highly Sensitised Patients on the UK Kidney Transplant Waiting List: a Matched Cohort Analysis. *The Lancet* (2017) 389:727–34. doi:10.1016/s0140-6736(16)31595-1
21. Orandi BJ, Garonzik-Wang JM, Massie AB, Zachary AA, Montgomery JR, Van Arendonk KJ, et al. Quantifying the Risk of Incompatible Kidney Transplantation: a Multicenter Study. *Am J Transpl* (2014) 14:1573–80. doi:10.1111/ajt.12786
22. Keith DS, Vranic GM. Approach to the Highly Sensitized Kidney Transplant Candidate. *Clin J Am Soc Nephrol* (2016) 11:684–93. doi:10.2215/cjn.05930615
23. Maggiore U, Oberbauer R, Pascual J, Viklicky O, Dudley C, Budde K, et al. EDTA-DESCARTES Working Group. Strategies to Increase the Donor Pool and Access to Kidney Transplantation: an International Perspective. *Nephrol Dial Transplant* (2015) 30:217–22. doi:10.1093/ndt/gfu212
24. Orandi BJ, Alachkar N, Kraus ES, Naqvi F, Lonze BE, Lees L, et al. Presentation and Outcomes of C4d-Negative Antibody-Mediated Rejection after Kidney Transplantation. *Am J Transplant* (2016) 16:213–20. doi:10.1111/ajt.13434
25. Jackson KR, Covarrubias K, Holscher CM, Luo X, Chen J, Massie AB, et al. The National Landscape of Deceased Donor Kidney Transplantation for the Highly Sensitized: Transplant Rates, Waitlist Mortality, and Posttransplant Survival under KAS. *Am J Transpl* (2019) 19:1129–38. doi:10.1111/ajt.15149
26. Heidt S, Haasnoot GW, Claas FHJ. How the Definition of Acceptable Antigens and Epitope Analysis Can Facilitate Transplantation of Highly Sensitized Patients with Excellent Long-Term Graft Survival. *Curr Opin Organ Transpl* (2018) 23:493–9. doi:10.1097/mot.0000000000000545
27. Hourmant M, Jacquelinet C, Antoine C, Hiesse C. Les nouvelles règles de répartition des greffons rénaux en France. *Néphrologie & Thérapeutique* (2005) 1:7–13. doi:10.1016/j.nephro.2005.01.002
28. Mumford L, Fuggle SV, Martorell J, Slavcev A, Iniotaki A, Haasnoot GW, et al. A Europe Wide Acceptable Mismatch Program Will Enable Transplantation of Long Waiting Highly Sensitised Patients with a Compatible Donor. *Transpl Immunol* (2021) 64:101354. doi:10.1016/j.trim.2020.101354
29. Schwaiger E, Eskandary F, Kozakowski N, Bond G, Kikić Ž, Yoo D, et al. Deceased Donor Kidney Transplantation across Donor-specific Antibody Barriers: Predictors of Antibody-Mediated Rejection. *Nephrol Dial Transpl* (2016) 31:1342–51. doi:10.1093/ndt/gfw027
30. Amrouche L, Aubert O, Suberbielle C, Rabant M, Van Huyen J-PD, Martinez F, et al. Long-term Outcomes of Kidney Transplantation in Patients with High Levels of Preformed DSA: the Necker High-Risk Transplant Program. *Transplantation* (2017) 101:2440–8. doi:10.1097/tp.0000000000001650
31. Huang E, Jordan SC. Immunoglobulin G-Degrading Enzyme of *Streptococcus Pyogenes(ides)*, Desensitization, and the Kidney Allocation System: Complementary Approaches to Increase Transplantation in Highly HLA Desensitized Patients. *Clin J Am Soc Nephrol* (2018) 13:799–801. doi:10.2215/cjn.12031017
32. Orandi BJ, Luo X, King EA, Garonzik-Wang JM, Bae S, Montgomery RA, et al. Hospital Readmissions Following HLA-Incompatible Live Donor Kidney Transplantation: A Multi-center Study. *Am J Transpl* (2018) 18:650–8. doi:10.1111/ajt.14472
33. Gaston RS, Cecka JM, Kasiske BL, Fieberg AM, Leduc R, Cosio FC, et al. Evidence for Antibody-Mediated Injury as a Major Determinant of Late Kidney Allograft Failure. *Transplantation* (2010) 90:68–74. doi:10.1097/tp.0b013e3181e065de
34. Senev A, Coemans M, Lerut E, Van Sandt V, Daniëls L, Kuypers D, et al. Histological Picture of Antibody-mediated Rejection without Donor-specific anti-HLA Antibodies: Clinical Presentation and Implications for Outcome. *Am J Transpl* (2019) 19:763–80. doi:10.1111/ajt.15074
35. Becker JU, Seron D, Rabant M, Roufosse C, Naesens M. Evolution of the Definition of Rejection in Kidney Transplantation and Its Use as an Endpoint in Clinical Trials. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10141
36. Haas M, Loupy A, Lefaucheur C, Roufosse C, Glotz D, Seron D, et al. The Banff 2017 Kidney Meeting Report: Revised Diagnostic Criteria for Chronic Active T Cell-Mediated Rejection, Antibody-mediated Rejection, and Prospects for Integrative Endpoints for Next-generation Clinical Trials. *Am J Transpl* (2018) 18:293–307. doi:10.1111/ajt.14625
37. Loupy A, Haas M, Roufosse C, Naesens M, Adam B, Afrouzian M, et al. The Banff 2019 Kidney Meeting Report (I): Updates on and Clarification of Criteria for T Cell- and Antibody-mediated Rejection. *Am J Transpl* (2020) 20:2318–31. doi:10.1111/ajt.15898
38. Kamburova EG, Wisse BW, Joosten I, Allebes WA, van der Meer A, Hilbrands LB, et al. Differential Effects of Donor-specific HLA Antibodies in Living versus Deceased Donor Transplant. *Am J Transpl* (2018) 18:2274–84. doi:10.1111/ajt.14709
39. Opelz G, Daniel V, Naujokat C, Fickenscher H, Döhler B. Effect of Cytomegalovirus Prophylaxis with Immunoglobulin or with Antiviral Drugs on post-transplant Non-hodgkin Lymphoma: a Multicentre Retrospective Analysis. *Lancet Oncol* (2007) 8:212–8. doi:10.1016/s1470-2045(07)70040-2
40. Williams RC, Opelz G, McGarvey CJ, Weil EJ, Chakkera HA. The Risk of Transplant Failure with HLA Mismatch in First Adult Kidney Allografts from Deceased Donors. *Transplantation* (2016) 100:1094–102. doi:10.1097/tp.0000000000001115
41. Duquesnoy RJ. A Structurally Based Approach to Determine HLA Compatibility at the Humoral Immune Level. *Hum Immunol* (2006) 67:847–62. doi:10.1016/j.humimm.2006.08.001
42. Tambur AR, Claas FHJ. HLA Epitopes as Viewed by Antibodies: what Is it All about? *Am J Transpl* (2015) 15:1148–54. doi:10.1111/ajt.13192
43. Wiebe C, Pochinco D, Blydt-Hansen TD, Ho J, Birk PE, Karpinski M, et al. Class II HLA Epitope Matching-A Strategy to MinimizeDe NovoDonor-specific Antibody Development and Improve Outcomes. *Am J Transpl* (2013) 13:3114–22. doi:10.1111/ajt.12478
44. Wiebe C, Rush DN, Nevins TE, Birk PE, Blydt-Hansen T, Gibson IW, et al. Class II Eplet Mismatch Modulates Tacrolimus Trough Levels Required to Prevent Donor-specific Antibody Development. *J Am Soc Nephrol* (2017) 28:335–62. doi:10.1681/asn.2017030287
45. Snanoudj R, Kamar N, Cassuto E, Caillard S, Metzger M, Merville P, et al. Epitope Load Identifies Kidney Transplant Recipients at Risk of Allosensitization Following Minimization of Immunosuppression. *Kidney Int* (2019) 95:1471–85. doi:10.1016/j.kint.2018.12.029
46. Senev A, Coemans M, Lerut E, Van Sandt V, Kerkhofs J, Daniëls L, et al. Eplet Mismatch Load and De Novo Occurrence of Donor-specific Anti-HLA Antibodies, Rejection, and Graft Failure after Kidney Transplantation: An Observational Cohort Study. *J Am Soc Nephrol* (2020) 31:2193–204. doi:10.1681/asn.2020010019
47. Kosmoliaptsis V, Chaudhry AN, Sharples LD, Halsall DJ, Dafforn TR, Bradley JA, et al. Predicting HLA Class I Alloantigen Immunogenicity from the Number and Physiochemical Properties of Amino Acid Polymorphisms. *Transplantation* (2009) 88:791–8. doi:10.1097/tp.0b013e3181b4a9ff
48. Kosmoliaptsis V, Sharples LD, Chaudhry AN, Halsall DJ, Bradley JA, Taylor CJ. Predicting HLA Class II Alloantigen Immunogenicity from the Number and Physiochemical Properties of Amino Acid Polymorphisms. *Transplantation* (2011) 91:183–90. doi:10.1097/tp.0b013e3181ffff99
49. Niemann M, Lachmann N, Geneugelijk K, Spierings E. Computational Eurotransplant Kidney Allocation Simulations Demonstrate the Feasibility and Benefit of T-Cell Epitope Matching. *Plos Comput Biol* (2021) 17:e1009248. doi:10.1371/journal.pcbi.1009248
50. Sicard A, Ducreux S, Rabeyrin M, Couzi L, McGregor B, Badet L, et al. Detection of C3d-Binding Donor-specific Anti-HLA Antibodies at Diagnosis of Humoral Rejection Predicts Renal Graft Loss. *J Am Soc Nephrol* (2015) 26:457–67. doi:10.1681/asn.2013101144
51. Mohan S, Palanisamy A, Tsapepas D, Tanriover B, Crew RJ, Dube G, et al. Donor-specific Antibodies Adversely Affect Kidney Allograft Outcomes. *J Am Soc Nephrol* (2012) 23:2061–71. doi:10.1681/asn.2012070664
52. Schinstock CA, GandhiCheungpasitporn MW, Cheungpasitporn W, Mitema D, Prieto M, Dean P, et al. Kidney Transplant with Low Levels of DSA or Low Positive B-Flow Crossmatch: an Underappreciated Option for Highly

Sensitized Transplant Candidates. *Transplantation* (2017) 101:2429–39. doi:10.1097/tp.0000000000001619
53. Ziemann M, Altermann W, Angert K, Arns W, Bachmann A, Bakchoul T, et al. Preformed Donor-specific HLA Antibodies in Living and Deceased Donor Transplantation: a Multicenter Study. *Clin J Am Soc Nephrol* (2019) 14: 1056–66. doi:10.2215/cjn.13401118
54. Thaunat O. Humoral Immunity in Chronic Allograft Rejection: Puzzle Pieces Come Together. *Transpl Immunol* (2012) 26:101–6. doi:10.1016/j.trim.2011.11.003
55. Roelen DL, Doxiadis II, Claas FHJ. Detection and Clinical Relevance of Donor Specific HLA Antibodies: a Matter of Debate. *Transpl Int* (2012) 25:604–10. doi:10.1111/j.1432-2277.2012.01491.x
56. Visentin J, Bachelet T, Aubert O, Del Bello A, Martinez C, Jambon F, et al. Reassessment of the Clinical Impact of Preformed Donor-specific anti-HLA-Cw Antibodies in Kidney Transplantation. *Am J Transpl* (2020) 20:1365–74. doi:10.1111/ajt.15766
57. Couzi L, Araujo C, Guidicelli G, Bachelet T, Moreau K, Morel D, et al. Interpretation of Positive Flow Cytometric Crossmatch in the Era of the Single-Antigen Bead Assay. *Transplantation* (2011) 91:527–35. doi:10.1097/tp.0b013e31820794bb
58. Bertrand D, Farce F, Laurent C, Hamelin F, François A, Guerrot D, et al. Comparison of Two Luminex Single-Antigen Bead Flow Cytometry Assays for Detection of Donor-specific Antibodies after Renal Transplantation. *Transplantation* (2019) 103:597–603. doi:10.1097/tp.0000000000002351
59. Bachelet T, Visentin J, Guidicelli G, Merville P, Couzi L, Taupin J-L. Anti-HLA Donor-specific Antibodies Are Not Created Equally. Don't Forget the Flow. *Transpl Int* (2016) 29:508–10. doi:10.1111/tri.12745
60. Meneghini M, Melilli E, Martorell J, Revuelta I, Rigol-Monzó E, Manonelles A, et al. Combining Sensitive Crossmatch Assays with Donor/recipient Human Leukocyte Antigen Eplet Matching Predicts Living-Donor Kidney Transplant Outcome. *Kidney Int Rep* (2018) 3:926–38. doi:10.1016/j.ekir.2018.03.015
61. Adebiyi OO, Gralla J, Klem P, Freed B, Davis S, Wiseman AC, et al. Clinical Significance of Pretransplant Donor-specific Antibodies in the Setting of Negative Cell-Based Flow Cytometry Crossmatching in Kidney Transplant Recipients. *Am J Transpl* (2016) 16:3458–67. doi:10.1111/ajt.13848
62. Huber L, Lachmann N, Niemann M, Naik M, Liefeldt L, Glander P, et al. Pretransplant Virtual PRA and Long-Term Outcomes of Kidney Transplant Recipients. *Transpl Int* (2015) 28:710–9. doi:10.1111/tri.12533
63. Dragun D, Catar R, Philippe A. Non-HLA Antibodies against Endothelial Targets Bridging Allo- and Autoimmunity. *Kidney Int* (2016) 90:280–8. doi:10.1016/j.kint.2016.03.019
64. Jackson AM, Sigdel TK, Delville M, Hsieh S-C, Dai H, Bagnasco S, et al. Endothelial Cell Antibodies Associated with Novel Targets and Increased Rejection. *J Am Soc Nephrol* (2015) 26:1161–71. doi:10.1681/asn.2013121277
65. Reindl-Schwaighofer R, Heinzel A, Kainz A, van Setten J, Jelencsics K, Hu K, et al. Contribution of Non-HLA Incompatibility between Donor and Recipient to Kidney Allograft Survival: Genome-wide Analysis in a Prospective Cohort. *The Lancet* (2019) 393:910–7. doi:10.1016/s0140-6736(18)32473-5
66. Montero N, Farouk S, Gandolfini I, Crespo E, Jarque M, Meneghini M, et al. Pretransplant Donor-specific IFNγ ELISPOT as a Predictor of Graft Rejection: A Diagnostic Test Accuracy Meta-Analysis. *Transpl Direct* (2019) 5:e451. doi:10.1097/txd.0000000000000886
67. Heidt S, Hester J, Shankar S, Friend PJ, Wood KJ. B Cell Repopulation after Alemtuzumab Induction-Transient Increase in Transitional B Cells and Long-Term Dominance of Naïve B Cells. *Am J Transplant* (2012) 12:1784–92. doi:10.1111/j.1600-6143.2012.04012.x
68. Luque S, Lúcia M, Melilli E, Lefaucheur C, Crespo M, Loupy A, et al. Value of Monitoring Circulating Donor-Reactive Memory B Cells to Characterize Antibody-Mediated Rejection after Kidney Transplantation. *Am J Transpl* (2019) 19:368–80. doi:10.1111/ajt.15055
69. Bestard O, Crespo E, Stein M, Lúcia M, Roelen DL, de Vaal YJ, et al. Cross-Validation of IFN-γ Elispot Assay for Measuring Alloreactive Memory/Effector T Cell Responses in Renal Transplant Recipients. *Am J Transpl* (2013) 13:1880–90. doi:10.1111/ajt.12285
70. Ashoor I, Najafian N, Korin Y, Reed EF, Mohanakumar T, Ikle D, et al. Standardization and Cross Validation of Alloreactive IFNγ ELISPOT Assays within the Clinical Trials in Organ Transplantation Consortium. *Am J Transpl* (2013) 13:1871–9. doi:10.1111/ajt.12286
71. Bestard O, Meneghini M, Crespo E, Bemelman F, Koch M, Volk HD, et al. Preformed T Cell Alloimmunity and HLA Eplet Mismatch to Guide Immunosuppression Minimization with Tacrolimus Monotherapy in Kidney Transplantation: Results of the CELLIMIN Trial. *Am J Transpl* (2021) 21:2833–45. doi:10.1111/ajt.16563
72. Lebranchu Y, Baan C, Biancone L, Legendre C, Morales JM, Naesens M, et al. Pretransplant Identification of Acute Rejection Risk Following Kidney Transplantation. *Transpl Int* (2014) 27:129–38. doi:10.1111/tri.12205

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10137

# Rationale for Surrogate Endpoints and Conditional Marketing Authorization of New Therapies for Kidney Transplantation

*Maarten Naesens[1]*, Alexandre Loupy[2], Luuk Hilbrands[3], Rainer Oberbauer[4], Maria Irene Bellini[5], Denis Glotz[6], Josep Grinyó[7], Uwe Heemann[8], Ina Jochmans[9], Liset Pengel[10], Marlies Reinders[11], Stefan Schneeberger[12] and Klemens Budde[13]*

[1]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium, [2]Paris Translational Research Center for Organ Transplantation, Hôpital Necker, Paris, France, [3]Department of Nephrology, Radboud University Medical Center, Nijmegen, Netherlands, [4]Department of Nephrology and Dialysis, Medical University of Vienna, Vienna, Austria, [5]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy, [6]Paris Translational Research Center for Organ Transplantation, Hôpital Saint Louis, Paris, France, [7]University of Barcelona, Barcelona, Spain, [8]Department of Nephrology, Technical University of Munich, Munich, Germany, [9]Transplantation Research Group, Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium, [10]Centre for Evidence in Transplantation, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom, [11]Erasmus MC Transplant Institute, Department of Internal Medicine, University Medical Center Rotterdam, Rotterdam, Netherlands, [12]Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria, [13]Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany

Conditional marketing authorization (CMA) facilitates timely access to new drugs for illnesses with unmet clinical needs, such as late graft failure after kidney transplantation. Late graft failure remains a serious, burdensome, and life-threatening condition for recipients. This article has been developed from content prepared by members of a working group within the European Society for Organ Transplantation (ESOT) for a Broad Scientific Advice request, submitted by ESOT to the European Medicines Agency (EMA), and reviewed by the EMA in 2020. The article presents the rationale for using surrogate endpoints in clinical trials aiming at improving late graft failure rates, to enable novel kidney transplantation therapies to be considered for CMA and improve access to medicines. The paper also provides background data to illustrate the relationship between primary and surrogate endpoints. Developing surrogate endpoints and a CMA strategy could be particularly beneficial for studies where the use of primary endpoints would yield insufficient statistical power or insufficient indication of long-term benefit following transplantation.

**Keywords: mortality, late graft failure, unmet medical need, morbidity, re-transplantation, clinical studies**

OPEN ACCESS

***Correspondence:***
*Maarten Naesens*
*maarten.naesens@kuleuven.be*

***Received:*** *21 October 2021*
***Accepted:*** *10 March 2022*
***Published:*** *20 May 2022*

***Citation:***
*Naesens M, Loupy A, Hilbrands L, Oberbauer R, Bellini MI, Glotz D, Grinyó J, Heemann U, Jochmans I, Pengel L, Reinders M, Schneeberger S and Budde K (2022) Rationale for Surrogate Endpoints and Conditional Marketing Authorization of New Therapies for Kidney Transplantation. Transpl Int 35:10137. doi: 10.3389/ti.2022.10137*

## INTRODUCTION

The guideline CHMP/EWP/263148/06 of the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP), issued in 2008, identifies the primary composite endpoint for clinical trials in organ transplantation as recipient death, graft failure, biopsy-confirmed acute rejection, and graft (dys)function (1). Based on this composite endpoint, specific immunosuppressive drugs have received full (standard) marketing authorization for

transplantation. However, CHMP/EWP/263148/06 does not mention any opportunities for other novel drugs to proceed to conditional marketing authorization (CMA), such as drugs that aim to improve long-term outcomes after kidney transplantation. This represents an area of considerable unmet medical need and restricts the development of novel treatments.

The present article proposes the rationale for surrogate endpoints for CMA, for novel kidney transplantation therapies; the paper also provides background data that illustrate the relationship between surrogate and primary endpoints, to support full marketing authorization.

CMA applications based on clinical trials using surrogate endpoints should not replace full marketing authorization applications based on studies using accepted primary endpoints. As discussed elsewhere in this Special Issue, graft rejection is acceptable as a primary endpoint for obtaining full marketing authorization by the EMA, because graft rejection is considered directly clinically meaningful, requiring therapies for rejection (2–4). Kidney function (incidence of end-stage renal disease, proportional decrease in eGFR, and annual decrease in eGFR—slope) is also well accepted by the EMA/CHMP as a primary endpoint to assess efficacy of medicinal products to slow progression of chronic renal insufficiency in chronic kidney disease. CHMP/EMA confirmed that this reasoning can be adopted for trials of kidney transplantation (5).

Rather, the CMA strategy and surrogate endpoints are suggested for studies where use of the accepted primary endpoints would yield insufficient statistical power or insufficient indication of long-term benefit. Applied to novel immunosuppressive agents, long-term benefit for kidney transplantation would equal decreased rates of late graft failure. It is therefore also important to have a very clear definition of late graft failure.

Here, we discuss the definition of late graft failure, and the rationale to consider late graft failure as a disease with unmet clinical need, allowing for CMA applications for novel therapies aimed at improving long-term kidney transplant outcomes. Endpoints that could be considered as surrogates for late graft failure are discussed separately in this Special Issue (6).

## DEFINITION OF LATE GRAFT FAILURE

In discussions relating to the present article, we defined overall (all-cause) graft failure as a composite of two important primary endpoints: loss of graft function (i.e., return to dialysis or pre-emptive re-transplantation), and recipient death with a functioning graft.

We consider that using 1 year post transplantation as the border between early and late graft failure reflects current clinical research standards and epidemiological data. These illustrate a fundamental difference in general improvement of graft outcome within and beyond 1 year after transplantation (7).

In addition, a 1-year threshold for the definition of late graft failure could be appropriate, given that research standards usually consider primary endpoints at 6 months to 1 year following transplantation. This was the case for pivotal trials that supported the approval of immunosuppressive drugs (reviewed in (8)). The 1-year threshold for early versus late graft failure also reflects evidence that short-term graft outcomes (i.e, failure within the first year) improve over time (7); this was not the case for long-term graft failure, which was defined as any failure from 1 year post-transplant (7). In addition, in the Collaborative Transplant Study European data analyses (9), the 1-year graft survival rate improved considerably between 1986 and 1999, but no noteworthy improvement was seen for graft survival beyond the first year after transplantation. Lastly, there are relevant differences in the reasons for graft loss in different periods after transplantation; it is not the purpose of the present paper to discuss them (10).

## RATIONALE FOR CMA APPLICATIONS FOR LATE GRAFT FAILURE

The European Medicines Agency (EMA)-initiated concept of CMA (11) is an important tool for ensuring timely access to medicines in areas of unmet medical need. For CMA application, medicines for human use are eligible if they belong to at least one of the following three categories:

- Aimed at treating, preventing, or diagnosing seriously debilitating or life-threatening diseases
- Intended for use in emergency situations (less-comprehensive pharmaceutical and non-clinical data may also be accepted)
- Designated as orphan medicines, i.e., for the diagnosis, prevention, or treatment of a life-threatening or chronically debilitating condition that is rare (affecting <5 in 10,000 people in the European Union [EU]).

### Late Graft Failure: Seriously Debilitating, Life-Threatening

In kidney transplant patient populations, late graft failure is a common, seriously debilitating, and life-threatening condition; no specific measures are available for its prevention. Immunosuppressive drugs were primarily approved for prevention of early acute rejection, with limited impact on (late) graft failure (8). In Europe, death-censored graft failure rates (censoring for death with a functioning graft) beyond the first year post-transplantation have shown some improvement since the late 1980s (7, 9). However, ~5% of grafts are still lost annually after the first year, including loss due to recipient death (7, 12, 13). On this basis alone, medicines that aim to prevent late kidney graft failure could be proposed for CMA.

Several aspects make late kidney graft failure a serious condition for which there is an unmet medical need. First, there is the requirement for dialysis reinitiation, which carries a heightened risk of mortality, comorbidities, and impaired health-related quality of life. Second, there is a high risk of human leukocyte antigen antibody (HLA) sensitization, which is associated with prolonged waiting time for repeat

transplantation and further increased risk of dialysis complications. Third, increased risk of graft failure is observed after re-transplantation, which is related to heightened risk of antibody-mediated rejection (AMR) because of preformed antibodies against the first donor kidney (13, 14). In addition, increased morbidity and inferior outcomes after re-transplantation can result from diverse complications such as long waiting times, increased doses of immunosuppressive therapy, increased risk of infections and malignancies, high rates of acute rejection, and delayed graft function. Kidney graft failure is also associated with increasing the average waiting time for transplantation, due to relisting (15).

As of December 31, 2019, at the time ESOT was discussing this issue, ~55,000 patients were on the transplantation waiting list in Europe (16), the vast majority of whom required kidney transplantation. Although ~16% of transplantations performed in 21 European countries were re-transplants (9), data from Eurotransplant (which includes a different spread of countries) show that >20% of patients on the kidney waiting list required re-transplantation after failure of a prior graft (17). Longer waiting time on dialysis is an independent risk factor for death (18), and a considerable proportion of patients with graft failure die while waiting for re-transplantation. For example, in 2019, ~10% of persons on the active Eurotransplant kidney waiting list were removed because they died or became unfit for transplantation (19).

While increasing longevity of kidney grafts could decrease the need for re-transplantation, importantly, the >20% of patients waitlisted for re-transplantation on Eurotransplant databases represents only those who are eligible for such procedures. Among European and US patients who experienced death-censored graft failure, 48% were waitlisted (median time 7.7 months) and 61% had HLA antibodies; most of the sensitized patients were not relisted for transplantation and remained on dialysis until death (20). A publication from Charité Hospital in Berlin found that between 1997 and 2017, 267 graft losses occurred in 254 patients, resulting in 117 (43.8%) relistings (21), of whom only 42 (35.9%) patients received a second transplant. At 5 years after graft loss, of the 254 patients, 49% had died, 27% were relisted, 14% were on dialysis and not relisted, and only 11% were re-transplanted (15).

Several studies demonstrate an increased mortality risk for patients who experience graft loss, compared with those with continued function (22–24) or those yet to receive a transplant (25). A study using competing-risk analysis confirmed a significantly increased all-cause mortality rate in patients relisted after graft failure compared with those awaiting a first transplant (16% vs. 11%; $p = 0.033$), with most deaths happening within 3 years of relisting (26). Prior transplant failure was associated with a 1.5-fold increased risk of mortality (95% confidence interval [CI] 1.01–2.2) (26).

However, a comparison of patients listed for first versus repeat transplantation does not account for the excess mortality rate seen in those who remain on lifelong dialysis after graft failure. Given that patients listed for re-transplantation are a selected population deemed capable of receiving another graft, it seems likely that those who are not relisted (primarily because of comorbidity and unacceptable risk) will have worse outcomes on dialysis. In addition, none of these analyses considers the burdens of returning to dialysis after failed transplantation, such as the costs associated with treatment (27), decreased ability to work and participate in society (28), and the psychological impact of returning to dialysis (29–31) (see also article by Tong et al. on patient reported outcome measures, in this Special Issue (32)).

### Late Graft Failure: An Orphan Indication?

In addition, late graft failure could be considered as an orphan indication, when its occurrence is calculated in absolute terms with the general population as reference. A hypothetical steady-state situation, where the same number of grafts are failing as are being transplanted, would result in ~21,000 graft losses per 512 million inhabitants in the EU, equivalent to four graft losses per 100,000 people, per year. This may fulfill the definition of an orphan indication and would do so even if twice as many graft losses were to occur.

## LATE GRAFT FAILURE: AN UNMET CLINICAL NEED

### Death With a Functioning Graft

Death of the recipient with a functioning graft is the most important reason for graft loss, and is usually a primary safety endpoint in studies of interventions that aim to prolong kidney transplant function. The main causes of death with a functioning graft are cardiovascular disease (CVD) and over-immunosuppression resulting in adverse events such as malignancy or infection (33–35). The fact that over-immunosuppression can cause death is obvious. Importantly, the relatively common side effects of immunosuppressants (e.g., diabetes mellitus, hypertension, altered lipid profile, and nephrotoxicity leading to low glomerular filtration rate) can also increase CVD risk (35). Graft function can also directly impact CVD risk and mortality, which provides further evidence for the pivotal role of good kidney function in both graft and patient survival (36, 37). The negative impact of poor kidney function on mortality (and CVD mortality in particular) is also seen in the general population (38, 39).

### Return to Dialysis/Re-Transplantation

Relative contributions of different pathological processes to graft failure have been evaluated (10, 33, 34, 40–42). Progression of fibrosis and accumulation of extracellular matrix i.e., interstitial fibrosis and tubular atrophy, IFTA) are key causes of graft loss. Fibrosis is thought to be mainly the consequence of nephron loss, and as aging is inevitably associated with a declining number of functioning nephrons, the quantity of nephrons might already be greatly reduced in grafts from marginal donors. After transplantation, nephrons can also be injured by immunological processes and/or other mechanisms (**Figure 1**) (43).

Increasing evidence suggests a continuous alloimmune response to the donor graft, despite modern immunosuppression, unrelated to the patient's level of adherence to immunotherapy. The incidence

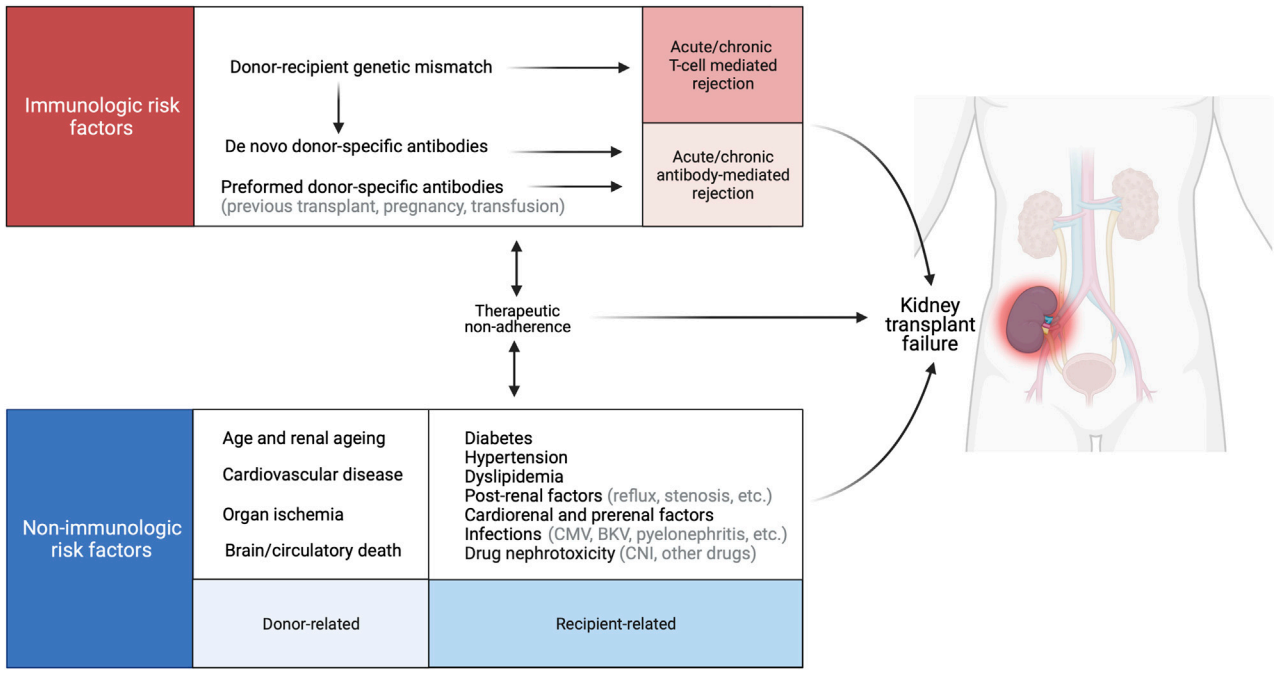


**FIGURE 1 |** Causes of late allograft loss. Late graft failure is a multifactorial process that involves immunological factors related to the donor–recipient genetic mismatch, and nonimmunological factors that contribute to acute and chronic lesion development. BKV, BK virus; CMV, cytomegalovirus, CNI, calcineurin inhibitor. Created with BioRender.com.

of acute cellular (i.e., T cell-mediated [TCMR]) rejection in the early months after transplantation is ~10% and rarely leads to immediate graft loss if treated appropriately, but TCMR is also an important, relevant, risk factor for long-term graft loss (10). Chronic TCMR has been described as a pathological entity and seems associated with impaired outcome, but its true prevalence and importance remain poorly defined (4, 44). By contrast, AMR diagnosis—and individual parameters of AMR—clearly show detrimental long-term effects on the graft (3, 10). B cells play key roles in AMR as antibody-producing cells and antigen-presenting cells for T cells with indirect allospecificity (12, 45). Poor adherence to medication is a major contributor to AMR development (10), highlighting that behavioral and social factors have important immunological consequences (43, 46). Poor adherence to complex medication regimens is common: it is estimated that up to 25% of patients have some degree of nonadherence, with severe nonadherence recognized as being a major contributor to late graft failure (10, 47). Poor adherence is associated with donor-specific antibody (DSA) development and poor control of metabolic factors (46).

As histologic studies show that progressive fibrosis is a major cause of late graft loss, and because calcineurin inhibitors (CNIs) are known to cause fibrosis, it was proposed that late graft loss might be partly attributable to CNI nephrotoxicity (10, 48), causing nephron injury and ultimately nephron loss with striped fibrosis. Studies have tested the hypothesis that minimizing the CNI dose, or avoiding these agents altogether, might improve long-term graft survival rates. Although some research suggested that avoiding CNIs did not cause safety issues and was associated with improvement in renal function over time, others indicated increased acute TCMR and DSA development in patients on CNI-sparing or CNI-free regimens and minimal, if any, improvement in renal function (49, 50). Thus, our understanding of the relative contribution of CNIs as the main cause of late kidney graft loss has evolved, and we recognize that competing risks (e.g., increased rate of rejection, or DSA development) might limit the success of CNI-sparing regimens.

After alloimmune-mediated injury, recurrence of native kidney disease in the transplanted organ is another common cause of graft loss (10, 51). Some native kidney diseases (e.g., focal segmental glomerulosclerosis or diseases associated with inherited complement defects) recur frequently, often early after transplantation and with poor ensuing graft survival. Although all kidney diseases are capable of recurrence, most do not strongly affect graft survival in the early years following transplantation. Of note, an elevated risk of late graft loss was observed in patients with recurring glomerulonephritis (12).

Nonimmunologic factors that contribute to post-transplantation nephron damage include brain death of the donor, poor donor management, and cold and warm ischemia times (52–54); delayed graft function (55); and infections (e.g., polyomavirus [BKV], cytomegalovirus, pyelonephritis) (10, 34). Kidney transplant recipients also usually have a high burden of comorbidities, some caused by chronic uremia before and during dialysis. Contributions of some modifiable CVD risk factors to the progression of native kidney disease have been demonstrated

unequivocally, but their effect on graft survival remains unclear because interventional studies are scarce. In competing-risk analyses, smoking, systolic blood pressure, and hemoglobin concentration remain as independent predictors of graft failure or doubling of creatinine level (12). Standard immunosuppressive regimens increase the risk of diabetes and hyperlipidemia, which appear to accelerate graft rejection independently of the potential effects of lipids on the graft vasculature (12).

Other factors that contribute to graft failure are reflux nephropathy or obstruction due to ureteral stenosis (10). Finally, poor graft quality (e.g., graft having lower reserves because of older donor age or expanded criteria donors) with lower nephron mass transplanted is an important baseline risk factor for late graft failure, as described previously (10).

Clearly, late graft failure is often a multifactorial process: active/acute diseases are additive and coincide with cumulative chronic injury (10, 12, 34, 56, 57). This chronic injury can also have many causes, increasing the vulnerability of grafts to superimposed acute injury. Acute and chronic factors (as described above) can injure the nephron; once this basic functional unit of the kidney is irreversibly damaged, it cannot be replaced, and renal function deteriorates. Hyperfiltration and glomerular hypertension of the remaining nephrons can lead to a vicious circle, with further reduction in functioning nephrons, as seen in native kidney disease. Although late graft failure is a heterogeneous condition, the underlying disease processes often share a common clinical pathway of declining kidney graft function (indicated by a declining glomerular filtration rate) and/or increasing proteinuria, with a rise in chronic histological injury and fibrosis.

Several studies highlight the importance of progressive fibrosis as a key pathway to graft failure and a target for intervention, independent of the recognized role of late AMR in graft failure (42, 44). Biopsies late after transplantation are particularly dominated by nonspecific chronic lesions and IFTA without displaying concomitant inflammation (44). Beyond 5–10 years after transplantation, failures become increasingly biased toward IFTA, which therefore represents a key finding among identifying factors involved in late graft failure. It is precisely these late failures that have proven so resistant to advances in transplantation practice (7, 9). However, underlying causes of IFTA and progressive nephron loss remain poorly understood: the histopathologic picture is complicated by issues including rejection phenomena and chronic CNI nephrotoxicity, together with under-investigated but clearly detrimental factors such as aging, viral infections, reflux, and pyelonephritis.

Progressive IFTA in the absence of inflammatory disease is a process once known as "mysterious dysregulated fibrosis" (40, 58). New insights have illuminated this process, which can involve epigenetic mechanisms, resulting in constitutive fibroblast activation (59), drug nephrotoxicity (60, 61) and other pathophysiological aspects (e.g., oxidative stress or innate immune activation (62)). Therapies directed toward progressive IFTA, which are emerging in the management of native kidney disease, should have some value after transplantation (62).

### Unmet Needs: Interventions to Improve Late Graft Failure

Current immunosuppressive agents were approved for marketing based on studies with follow-up periods of <1 year. The approval of drugs that improved these short-term outcomes was based on research focusing on TCMR inhibition, which led to an important decline in early graft failure rates (7, 9) but did not substantially benefit long-term outcomes.

The impact of older immunosuppressive agents (e.g., cyclosporine) is not limited to short-term endpoints, however. Studies with ≥5-years follow-up periods, including cyclosporine withdrawal regimens, have demonstrated the effect of immunosuppressive drugs on long-term graft outcomes (63, 64). This suggests that different competing risks exist at different time points following transplantation. In addition, studies with tacrolimus have illustrated improved long-term outcomes compared with cyclosporine (65).

Very few randomized controlled trials (RCTs) have evaluated newer immunosuppressive agents (e.g., mTOR inhibitors, interleukin-2 receptor blockade, belatacept) with long-term graft survival as an endpoint. Extensions of the BENEFIT studies, reported at 7 and 10 years post transplantation (66, 67), demonstrated significantly lower risk of death or graft failure in the belatacept-treated group versus the cyclosporine-treated group, but only in standard criteria donor transplantations (67). Belatacept-treated patients had better outcomes despite having experienced more severe rejections (mainly TCMR) in the first year (66, 67), similar to findings of a study of early CNI withdrawal that included extensive follow-up (68). These studies clearly demonstrate the dissociation between TCMR and long-term outcome, suggesting that competing risks (e.g., cyclosporine toxicity, differences in metabolic profile, *de novo* DSA development) are more important than TCMR for long-term transplantation success.

Other studies had extended follow-up (beyond 1 year) after transplantation, comparing regimens of immunosuppressive agents that were approved based on short-term data. Although graft function sometimes improved over time, this did not reduce the rates of long-term graft failure (68, 69). Sometimes, worsening graft function and long-term graft survival rates were observed for the innovative regimen (70), which supports the hypothesis that long-term graft survival is affected by different competing risks at different time points. The complex reasons for graft loss (10), and the paucity of RCTs investigating the translation of short-term results into long-term survival benefits, highlight the difficulties in powering such trials sufficiently. Interpretation of long-term follow-up data is also confounded by frequent conversions to new, different immunosuppressive regimens.

## SURROGATE ENDPOINTS FOR CMA APPLICATIONS FOR LATE GRAFT FAILURE

If CMA applications for novel drugs aiming at preventing or treating late graft failure are admissible to the EMA, the next discussion relates to the choice of the endpoints to be used for the required clinical trials. Graft failure is a highly relevant hard endpoint in clinical studies, but it is a late endpoint. This hampers

the feasibility of using graft failure as an endpoint in clinical trials that aim at improving late graft failure rates.

Surrogates for late graft failure are therefore needed but require robust definitions. A good surrogate endpoint should fulfil four criteria: 1) The disease process is sufficiently understood; 2) The surrogate endpoint has biologic plausibility; 3) The strength of the consistency supports the relationship between the surrogate marker and outcome; 4) Treatment effects on the surrogate endpoint predict treatment effects on the clinical outcome of interest.

Kidney graft function and combined functional markers, donor-specific HLA antibodies and composite scores could be considered as surrogate endpoints, but do not fulfill all these criteria. For a detailed discussion on the potential acceptability of these surrogate endpoints for late graft failure, we refer to another manuscript in this Special Issue (6).

## FROM CONDITIONAL TO FULL MARKETING AUTHORIZATION

After successful application for CMA of a product aimed at improving long-term graft survival, conversion to full marketing authorization is necessary, based on a post-marketing confirmatory commitment.

ESOT sees different options for this conversion of CMA to full marketing authorization. For example, applicants could consider requests for full marketing authorization based on long-term registration studies with accepted primary endpoints relating to graft rejection (2–4) function (5) and/or graft failure. Applicants could also consider requesting full marketing authorization based on comprehensive high-quality evidence from open-label study data, comparing findings to appropriate historic controls.

Alternatively, applicants could base the comprehensive evidence for full marketing authorization requests on good-quality data from registration studies, utilizing real-world data. Indeed, the EMA has already considered data from two other registries suitable for their decision-making processes: the European Cystic Fibrosis Society patient registry and the Cellular Therapy module of the European Blood and Marrow Transplant registry. The EMA Patient Registries Initiative (71) offers guidance on this topic. Of note, ESOT emphasizes that currently no European registries in kidney transplantation could be used as basis for requesting full marketing authorization.

A final option could be to use data from a qualified surrogate endpoint as a source of comprehensive evidence for a full marketing authorization request. Although CHMP/EMA has suggested to initiate a formal Qualification of Novel Methodologies procedure for e.g. the finalized iBox model (69) as a surrogate marker, this qualification is not yet achieved. The status and path toward formal qualification of composite scores as potential surrogate endpoints is discussed separately in this Special Issue (6).

Each of the above options for post-marketing commitments seems unsatisfactory at present, in the field of kidney transplantation. This may hamper the current admissibility of CMA applications for therapies aiming at reducing the incidence and burden of late kidney transplant failure. The results of formal qualification procedures are eagerly awaited and will hopefully change the landscape in future.

## CONCLUSION

- Late graft failure (loss of graft function >1 year post transplant) is a condition with unmet medical need. Therefore, CMA should be considered for interventions that demonstrate potential benefits:
  - Late graft failure is a seriously debilitating, life-threatening disease for which no specific preventive or treatment options are available.
  - CMA of therapies aimed at preventing late graft failure could be based on trials that show benefit on a validated surrogate endpoint for graft failure.
- For drugs aimed at reducing late graft failure, applying for CMA could be considered.
  - CMA procedures facilitate timely access to new therapies.
  - Confirmatory post-marketing commitments will be needed to convert CMA to full marketing authorization.

### Scientific Advice From the Committee for Medicinal Products for Human Use (CHMP) of the European Medicines Agency (EMA) Regarding These Conclusions

- The CHMP agreed that improving long-term outcome after kidney transplantation is an area of unmet medical need; arguments for orphan designation of late graft failure were not followed.
- Should a novel therapy be proposed for CMA, the product will need to fulfil all of the following four criteria at the time CMA is considered: 1) positive benefit/risk balance; 2) it is likely that the applicant will be able to provide comprehensive data later; 3) unmet medical need is fulfilled; and 4) the benefit to public health of the medicinal product's immediate availability on the market outweighs the risk due to need for further data.
- Criteria for CMA will be reviewed for specific data submitted; CMA cannot be granted *a priori* for any given product or indication.

## AUTHOR CONTRIBUTIONS

This article is one of several developed from the Broad Scientific Advice request, submitted to the EMA/CHMP by ESOT in 2020. For the present article, the working group on surrogate endpoints in kidney transplantation developed the ESOT position on the core question "Does CHMP agree that long-term outcome after kidney transplantation is an area of unmet medical need, for which conditional marketing authorization procedures should be considered, to facilitate timely access to new therapies? If so, does CHMP agree with the proposed surrogate endpoints for clinical trials for therapies requiring conditional marketing authorization?"

The Centre for Evidence in Transplantation provided support with specific data extraction requests: these literature searches formed the basis of evidence used in the advice request and the present article. Input into the working group's output was provided from all ESOT members involved in the advice request process.

The present article was adapted by MN from the final Broad Scientific Advice request submission (June 2020), documentation from the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020). Drafts of the article were circulated to all co-authors for review and approval before submission.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

LH reports speaker fees from Astellas, consultancy and research support from Chiesi, consultancy for Novartis, and research support from Sandoz. RO has received grants/research support from Amgen, Astellas, and Chiesi; and speakers' bureaux/honoraria from Amgen, Astellas, Chiesi, Hansa, Neovii, Novartis, and Teva. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva. DG reports consultancy agreements and scientific boards for AstraZeneca, BMS, Hansa, and Sanofi. UH has received grants/research support from Baxter, Chiesi, and Neovii; speakers' bureaux/honoraria from Chiesi and Hansa; and consulting fees from Astellas, Hansa, Neovii, Novartis, and Teva. IJ's institution has received speaker's fees from XVIVO Perfusion. MR has received lecture fees from Astellas and Chiesi; and research grant support (paid to institution) from Astellas and Chiesi for investigator-initiated studies. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva. KB has received honoraria and/or research funding from Alexion, Astellas, Bristol Myers Squibb, Chiesi, Fresenius, Hansa, Hexal, Merck, Novartis, Otsuka, Pfizer, Roche, Sandoz, Siemens, and Veloxis.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. European Medicines Agency. Clinical Investigation of Immunosuppressants for Solid Organ Transplantation (2008). CHMP/EWP/263148/06. Available at: https://www.ema.europa.eu/en/clinical-investigation-immunosuppressants-solid-organ-transplantation (Accessed February 8, 2022).
2. Becker JU, Seron D, Rabant M, Roufosse C, Naesens M. Evolution of the Definition of Rejection in Kidney Transplantation and Its Use as an Endpoint in Clinical Trials. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10141
3. Roufosse C, Becker JU, Rabant M, Seron D, Bellini MI, Böhmig G, et al. Proposed Definitions of Antibody-Mediated Rejection for Use as a Clinical Trial Endpoint in Kidney Transplantation. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10140
4. Seron D, Rabant M, Becker JU, Roufosse C, Bellini MI, Böhmig G, et al. Proposed Definitions of T Cell-Mediated Rejection and Tubulointerstitial Inflammation as Clinical Trial Endpoints in Kidney Transplantation. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10135
5. Hilbrands L, Budde K, Bellini MI, Diekmann F, Furian L, Grinyó J, et al. Allograft Function as Endpoint for Clinical Trials in Kidney Transplantation. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10139
6. Naesens M, Budde K, Hilbrands L, Oberbauer R, Bellini MI, Glotz D, et al. Surrogate Endpoints for Late Kidney Transplantation Failure. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10136
7. Lamb KE, Lodhi S, Meier-Kriesche H-U. Long-term Renal Allograft Survival in the United States: a Critical Reappraisal. *Am J Transpl* (2011) 11:450–62. doi:10.1111/j.1600-6143.2010.03283.x
8. Naesens M, Thaunat O. BENEFIT of Belatacept: Kidney Transplantation Moves Forward. *Nat Rev Nephrol* (2016) 12:261–2. doi:10.1038/nrneph.2016.34
9. Coemans M, Süsal C, Döhler B, Anglicheau D, Giral M, Bestard O, et al. Analyses of the Short- and Long-Term Graft Survival after Kidney Transplantation in Europe between 1986 and 2015. *Kidney Int* (2018) 94:964–73. doi:10.1016/j.kint.2018.05.018
10. Mayrdorfer M, Liefeldt L, Wu K, Rudolph B, Zhang Q, Friedersdorff F, et al. Exploring the Complexity of Death-Censored Kidney Allograft Failure. *J Am Soc Nephrol* (2021) 32:1513–26. doi:10.1681/asn.2020081215
11. European Medicines Agency. Conditional Marketing Authorization (2022). Available at: https://www.ema.europa.eu/en/human-regulatory/marketing-authorisation/conditional-marketing-authorisation (Accessed February 8, 2022).
12. Wekerle T, Segev D, Lechler R, Oberbauer R. Strategies for Long-Term Preservation of Kidney Graft Function. *The Lancet* (2017) 389:2152–62. doi:10.1016/s0140-6736(17)31283-7
13. Agence de la Biomedicine. Annual Report (2017). Available at: https://www.agence-biomedecine.fr/annexes/bilan2017/donnees/organes/06-rein/synthese.htm (Accessed February 8, 2022).
14. Süsal C, Döhler B, Opelz G. Presensitized Kidney Graft Recipients with HLA Class I and II Antibodies Are at Increased Risk for Graft Failure: a Collaborative Transplant Study Report. *Hum Immunol* (2009) 70:569–73. doi:10.1016/j.humimm.2009.04.013

15. Schrezenmeier E, Lehner LJ, Merkel M, Mayrdorfer M, Duettmann W, Naik MG, et al. What Happens after Graft Loss? A Large, Long-term, Single-center Observation. *Transpl Int* (2021) 34:732–42. doi:10.1111/tri.13834
16. DG SANTE, European Commission. Organs, Blood, Tissues & Cells in the EU (2022). Available at: https://ec.europa.eu/health//sites/health/files/blood_tissues_organs/docs/infographic_obtc_en.pdf (Accessed February 8, 2022).
17. Eurotransplant statistics library (2021). Available at: https://www.eurotransplant.org/statistics/statistics-library/ (Accessed February 8, 2022).
18. Haller MC, Kainz A, Baer H, Oberbauer R. Dialysis Vintage and Outcomes after Kidney Transplantation: a Retrospective Cohort Study. *Clin J Am Soc Nephrol* (2017) 12:122–30. doi:10.2215/cjn.04120416
19. Eurotransplant. (2019). Available at: https://www.eurotransplant.org/annual-report/annual-report-2019/ (Accessed February 8, 2022). Annual Report.
20. Rose C, Schaeffner E, Frei U, Gill J, Gill JS. A Lifetime of Allograft Function with Kidneys from Older Donors. *J Am Soc Nephrol* (2015) 26:2483–93. doi:10.1681/asn.2014080771
21. Halleck F, Merkel M, Bach P, Niemann M, Lachmann N, Lehner L, et al. Epitope Matching Predicts Immunization after Allograft Failure Causing Longer Waiting Time for Re-transplantation. ERA-EDTA 2019 Congress Abstracts Library; SaO017 (2022). Available at: https://atcmeetingabstracts.com/abstract/epitop-matching-predicts-immunization-in-kidney-transplant-recipients-with-allograft-failure-leading-to-worse-outcomes/ (Accessed February 8, 2022).
22. Kaplan B, Meier-Kriesche H-U. Death after Graft Loss: an Important Late Study Endpoint in Kidney Transplantation. *Am J Transpl* (2002) 2:970–4. doi:10.1034/j.1600-6143.2002.21015.x
23. Rao PS, Schaubel DE, Saran R. Impact of Graft Failure on Patient Survival on Dialysis: a Comparison of Transplant-Naive and post-graft Failure Mortality Rates. *Nephrol Dial Transplant* (2005) 20:387–91. doi:10.1093/ndt/gfh595
24. Rao PS, Schaubel DE, Jia X, Li S, Port FK, Saran R. Survival on Dialysis post-kidney Transplant Failure: Results from the Scientific Registry of Transplant Recipients. *Am J Kidney Dis* (2007) 49:294–300. doi:10.1053/j.ajkd.2006.11.022
25. Perl J, Zhang J, Gillespie B, Wikstrom B, Fort J, Hasegawa T, et al. Reduced Survival and Quality of Life Following Return to Dialysis after Transplant Failure: the Dialysis Outcomes and Practice Patterns Study. *Nephrol Dial Transplant* (2012) 27:4464–72. doi:10.1093/ndt/gfs386
26. Hernández D, Muriel A, Castro de la Nuez P, Alonso-titos J, Ruiz-Esteban P, Duarte A, et al. Survival in Southern European Patients Waitlisted for Kidney Transplant after Graft Failure: a Competing Risk Analysis. *PLoS One* (2018) 13:e0193091. doi:10.1371/journal.pone.0193091
27. Vanholder R, Davenport A, Hannedouche T, Kooman J, Kribben A, Lameire N, et al. Reimbursement of Dialysis: a Comparison of Seven Countries. *J Am Soc Nephrol* (2012) 23:1291–8. doi:10.1681/asn.2011111094
28. Erickson KF, Zhao B, Ho V, Winkelmayer WC. Employment Among Patients Starting Dialysis in the United States. *Clin J Am Soc Nephrol* (2018) 13:265–73. doi:10.2215/cjn.06470617
29. Palmer S, Vecchio M, Craig JC, Tonelli M, Johnson DW, Nicolucci A, et al. Prevalence of Depression in Chronic Kidney Disease: Systematic Review and Meta-Analysis of Observational Studies. *Kidney Int* (2013) 84:179–91. doi:10.1038/ki.2013.77
30. Tong A, Hanson CS, Chapman JR, Halleck F, Budde K, Josephson MA, et al. 'Suspended in a Paradox'-Patient Attitudes to Wait-Listing for Kidney Transplantation: Systematic Review and Thematic Synthesis of Qualitative Studies. *Transpl Int* (2015) 28:771–87. doi:10.1111/tri.12575
31. Tong A, Gill J, Budde K, Marson L, Reese PP, Rosenbloom D, et al. Toward Establishing Core Outcome Domains for Trials in Kidney Transplantation. *Transplantation* (2017) 101:1887–96. doi:10.1097/tp.0000000000001774
32. Tong A, Oberbauer R, Bellini MI, Budde K, Caskey FJ, Dobbels F, et al. Patient-Reported Outcomes as Endpoints in Clinical Trials of Kidney Transplantation Interventions. *Transpl Int* (2022) 35. doi:10.3389/ti.2022.10134
33. El-Zoghby ZM, Stegall MD, Lager DJ, Kremers WK, Amer H, Gloor JM, et al. Identifying Specific Causes of Kidney Allograft Loss. *Am J Transpl* (2009) 9:527–35. doi:10.1111/j.1600-6143.2008.02519.x
34. Van Loon E, Senev A, Lerut E, Coemans M, Callemeyn J, Van Keer JM, et al. Assessing the Complex Causes of Kidney Allograft Loss. *Transplantation* (2020) 104:2557–66. doi:10.1097/tp.0000000000003192
35. Liefeldt L, Budde K. Risk Factors for Cardiovascular Disease in Renal Transplant Recipients and Strategies to Minimize Risk. *Transpl Int* (2010) 23:1191–204. doi:10.1111/j.1432-2277.2010.01159.x
36. Fellstrom B, Jardine AG, Soveri I, Cole E, Neumayer H-H, Maes B, et al. Renal Dysfunction Is a strong and Independent Risk Factor for Mortality and Cardiovascular Complications in Renal Transplantation. *Am J Transpl* (2005) 5:1986–91. doi:10.1111/j.1600-6143.2005.00983.x
37. Soveri I, Snyder J, Holdaas H, Holme I, Jardine AG, L'Italien GJ, et al. The External Validation of the Cardiovascular Risk Equation for Renal Transplant Recipients. *Transplantation* (2013) 95:142–7. doi:10.1097/tp.0b013e31827722c9
38. Badve SV, Palmer SC, Hawley CM, Pascoe EM, Strippoli GFM, Johnson DW. Glomerular Filtration Rate Decline as a Surrogate End point in Kidney Disease Progression Trials. *Nephrol Dial Transpl* (2016) 31:1425–36. doi:10.1093/ndt/gfv269
39. Astor BC, Matsushita K, Matsushita K, Gansevoort RT, van der Velde M, Woodward M, et al. Lower Estimated Glomerular Filtration Rate and Higher Albuminuria Are Associated with Mortality and End-Stage Renal Disease. A Collaborative Meta-Analysis of Kidney Disease Population Cohorts. *Kidney Int* (2011) 79:1331–40. doi:10.1038/ki.2010.550
40. Sellarés J, De Freitas DG, Mengel M, Reeve J, Einecke G, Sis B, et al. Understanding the Causes of Kidney Transplant Failure: the Dominant Role of Antibody-Mediated Rejection and Nonadherence. *Am J Transpl* (2012) 12:388–99. doi:10.1111/j.1600-6143.2011.03840.x
41. Naesens M, Kuypers DRJ, De Vusser K, Evenepoel P, Claes K, Bammens B, et al. The Histology of Kidney Transplant Failure. *Transplantation* (2014) 98:427–35. doi:10.1097/tp.0000000000000183
42. Chand S, Atkinson D, Collins C, Briggs D, Ball S, Sharif A, et al. The Spectrum of Renal Allograft Failure. *PLoS One* (2016) 11:e0162278. doi:10.1371/journal.pone.0162278
43. Neuberger JM, Bechstein WO, Kuypers DR, Burra P, Citterio F, De Geest S, et al. Practical Recommendations for Long-Term Management of Modifiable Risks in Kidney and Liver Transplant Recipients: a Guidance Report and Clinical Checklist by the Consensus on Managing Modifiable Risk in Transplantation (COMMIT) Group. *Transplantation* (2017) 101(4S Suppl. 2):S1–56. doi:10.1097/TP.0000000000001651
44. Haas M, Loupy A, Lefaucheur C, Roufosse C, Glotz D, Seron D, et al. The Banff 2017 Kidney Meeting Report: Revised Diagnostic Criteria for Chronic Active T Cell-Mediated Rejection, Antibody-mediated Rejection, and Prospects for Integrative Endpoints for Next-generation Clinical Trials. *Am J Transpl* (2018) 18:293–307. doi:10.1111/ajt.14625
45. Loupy A, Lefaucheur C. Antibody-mediated Rejection of Solid-Organ Allografts. *N Engl J Med* (2018) 379:1150–60. doi:10.1056/nejmra1802677
46. Gokoel SRM, Gombert-Handoko KB, Zwart TC, van der Boog PJM, Moes DJAR, de Fijter JW. Medication Non-adherence after Kidney Transplantation: A Critical Appraisal and Systematic Review. *Transplant Rev* (2020) 34:100511. doi:10.1016/j.trre.2019.100511
47. Morath C, Opelz G, Zeier M, Süsal C. Prevention of Antibody-Mediated Kidney Transplant Rejection. *Transpl Int* (2012) 25:633–45. doi:10.1111/j.1432-2277.2012.01490.x
48. Jevnikar AM, Mannon RB. Late Kidney Allograft Loss: what We Know about it, and what We Can Do about it. *Clin J Am Soc Nephrol* (2008) 3(Suppl. 2):S56–67. doi:10.2215/CJN.03040707
49. Sharif A, Shabir S, Chand S, Cockwell P, Ball S, Borrows R. Meta-analysis of Calcineurin-Inhibitor-Sparing Regimens in Kidney Transplantation. *J Am Soc Nephrol* (2011) 22:2107–18. doi:10.1681/asn.2010111160
50. Lim WH, Eris J, Kanellis J, Pussell B, Wiid Z, Witcombe D, et al. A Systematic Review of Conversion from Calcineurin Inhibitor to Mammalian Target of Rapamycin Inhibitors for Maintenance Immunosuppression in Kidney Transplant Recipients. *Am J Transpl* (2014) 14:2106–19. doi:10.1111/ajt.12795
51. Briganti EM, Russ GR, McNeil JJ, Atkins RC, Chadban SJ. Risk of Renal Allograft Loss from Recurrent Glomerulonephritis. *N Engl J Med* (2002) 347:103–9. doi:10.1056/nejmoa013036
52. Summers DM, Johnson RJ, Allen J, Fuggle SV, Collett D, Watson CJ, et al. Analysis of Factors that Affect Outcome after Transplantation of Kidneys Donated after Cardiac Death in the UK: a Cohort Study. *The Lancet* (2010) 376:1303–11. doi:10.1016/s0140-6736(10)60827-6
53. Heylen L, Jochmans I, Samuel U, Tieken I, Naesens M, Pirenne J, et al. The Duration of Asystolic Ischemia Determines the Risk of Graft Failure after

Circulatory-Dead Donor Kidney Transplantation: a Eurotransplant Cohort Study. *Am J Transpl* (2018) 18:881–9. doi:10.1111/ajt.14526
54. Krishnan AR, Wong G, Chapman JR, Coates PT, Russ GR, Pleass H, et al. Prolonged Ischemic Time, Delayed Graft Function, and Graft and Patient Outcomes in Live Donor Kidney Transplant Recipients. *Am J Transpl* (2016) 16:2714–23. doi:10.1111/ajt.13817
55. Butala NM, Reese PP, Doshi MD, Parikh CR. Is Delayed Graft Function Causally Associated with Long-Term Outcomes after Kidney Transplantation? Instrumental Variable Analysis. *Transplantation* (2013) 95:1008–14. doi:10.1097/tp.0b013e3182855544
56. Legendre C, Canaud G, Martinez F. Factors Influencing Long-Term Outcome after Kidney Transplantation. *Transpl Int* (2014) 27:19–27. doi:10.1111/tri.12217
57. Van Loon E, Lerut E, Naesens M. The Time Dependency of Renal Allograft Histology. *Transpl Int* (2017) 30:1081–91. doi:10.1111/tri.13042
58. Naesens M, Lerut E. Looking Back to Evaluate the Causes of Graft Loss? A Response to Dr. Halloran and Dr. Sellares.*Am J Transpl* (2013) 13:1933–4. doi:10.1111/ajt.12316
59. Bechtel W, McGoohan S, Zeisberg EM, Müller GA, Kalbacher H, Salant DJ, et al. Methylation Determines Fibroblast Activation and Fibrogenesis in the Kidney. *Nat Med* (2010) 16:544–50. doi:10.1038/nm.2135
60. Chapman JR. Chronic Calcineurin Inhibitor Nephrotoxicity-Lest We Forget. *Am J Transpl* (2011) 11:693–7. doi:10.1111/j.1600-6143.2011.03504.x
61. Naesens M, Kuypers DRJ, Sarwal M. Calcineurin Inhibitor Nephrotoxicity. *Clin J Am Soc Nephrol* (2009) 4:481–508. doi:10.2215/cjn.04800908
62. Boor P, Floege J. Renal Allograft Fibrosis: Biology and Therapeutic Targets. *Am J Transpl* (2015) 15:863–86. doi:10.1111/ajt.13180
63. Gallagher M, Jardine M, Perkovic V, Cass A, McDonald S, Petrie J, et al. Cyclosporine Withdrawal Improves Long-Term Graft Survival in Renal Transplantation. *Transplantation* (2009) 87:1877–83. doi:10.1097/tp.0b013e3181a76823
64. Kreis H, Oberbauer R, Campistol JM, Mathew T, Daloze P, Schena FP, et al. Long-term Benefits with Sirolimus-Based Therapy after Early Cyclosporine Withdrawal. *J Am Soc Nephrol* (2004) 15:809–17. doi:10.1097/01.asn.0000113248.59077.76
65. Jurewicz WA. Tacrolimus versus Cyclosporin Immunosuppression: Long-Term Outcome in Renal Transplantation. *Nephrol Dial Transpl* (2003) 18(Suppl. 1):i7–11. doi:10.1093/ndt/gfg1028
66. Vincenti F, Blancho G, Durrbach A, Grannas G, Grinyó J, Meier-Kriesche H-U, et al. Ten-year Outcomes in a Randomized Phase II Study of Kidney Transplant Recipients Administered Belatacept 4-weekly or 8-weekly. *Am J Transpl* (2017) 17:3219–27. doi:10.1111/ajt.14452
67. Durrbach A, Pestana JM, Florman S, Del Carmen Rial M, Rostaing L, Kuypers D, et al. Long-term Outcomes in Belatacept- versus Cyclosporine-Treated Recipients of Extended Criteria Donor Kidneys: Final Results from BENEFIT-EXT, a Phase Iii Randomized Study. *Am J Transpl* (2016) 16:3192–201. doi:10.1111/ajt.13830
68. Budde K, Lehner F, Sommerer C, Reinke P, Arns W, Eisenberger U, et al. Five-year Outcomes in Kidney Transplant Patients Converted from Cyclosporine to Everolimus: the Randomized ZEUS Study. *Am J Transpl* (2014) 15:119–28. doi:10.1111/ajt.12952
69. Sommerer C, Duerr M, Witzke O, Lehner F, Arns W, Kliem V, et al. Five-year Outcomes in Kidney Transplant Patients Randomized to Everolimus with Cyclosporine Withdrawal or Low-Exposure Cyclosporine versus Standard Therapy. *Am J Transpl* (2018) 18: 2965–76. doi:10.1111/ajt.14897
70. Chhabra D, Skaro AI, Leventhal JR, Dalal P, Shah G, Wang E, et al. Long-term Kidney Allograft Function and Survival in Prednisone-free Regimens: Tacrolimus/mycophenolate Mofetil versus Tacrolimus/sirolimus. *Clin J Am Soc Nephrol* (2012) 7:504–12. doi:10.2215/cjn.06940711
71. EMA Human Regulatory: Patient Registries (2021). Available at: https://www.ema.europa.eu/en/human-regulatory/post-authorisation/patient-registries (Accessed February 8, 2022).

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10136

# Surrogate Endpoints for Late Kidney Transplantation Failure

*Maarten Naesens[1]*, Klemens Budde[2], Luuk Hilbrands[3], Rainer Oberbauer[4], Maria Irene Bellini[5], Denis Glotz[6], Josep Grinyó[7], Uwe Heemann[8], Ina Jochmans[1], Liset Pengel[9], Marlies Reinders[10], Stefan Schneeberger[11] and Alexandre Loupy[12]*

[1]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium, [2]Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany, [3]Department of Nephrology, Radboud University Medical Center, Nijmegen, Netherlands, [4]Department of Nephrology and Dialysis, Medical University of Vienna, Vienna, Austria, [5]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy, [6]Paris Translational Research Center for Organ Transplantation, Hôpital Saint Louis, Paris, France, [7]University of Barcelona, Barcelona, Spain, [8]Department of Nephrology, Technical University of Munich, Munich, Germany, [9]Centre for Evidence in Transplantation, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom, [10]Department of Internal Medicine, Erasmus MC Transplant Institute, University Medical Center Rotterdam, Rotterdam, Netherlands, [11]Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria, [12]Paris Translational Research Center for Organ Transplantation, Hôpital Necker, Paris, France

In kidney transplant recipients, late graft failure is often multifactorial. In addition, primary endpoints in kidney transplantation studies seek to demonstrate the short-term efficacy and safety of clinical interventions. Although such endpoints might demonstrate short-term improvement in specific aspects of graft function or incidence of rejection, such findings do not automatically translate into meaningful long-term graft survival benefits. Combining many factors into a well-validated model is therefore more likely to predict long-term outcome and better reflect the complexity of late graft failure than using single endpoints. If conditional marketing authorization could be considered for therapies that aim to improve long-term outcomes following kidney transplantation, then the surrogate endpoint for graft failure in clinical trial settings needs clearer definition. This Consensus Report considers the potential benefits and drawbacks of several candidate surrogate endpoints (including estimated glomerular filtration rate, proteinuria, histological lesions, and donor-specific anti-human leukocyte antigen antibodies) and composite scoring systems. The content was created from information prepared by a working group within the European Society for Organ Transplantation (ESOT). The group submitted a Broad Scientific Advice request to the European Medicines Agency (EMA), June 2020: the request focused on clinical trial design and endpoints in kidney transplantation. Following discussion and refinement, the EMA made final recommendations to ESOT in December 2020 regarding the potential to use surrogate endpoints in clinical studies that aim to improving late graft failure.

**Keywords: rejection, outcome, graft function, conditional marketing authorization, iBox**

OPEN ACCESS

***Correspondence:***
*Maarten Naesens*
*maarten.naesens@kuleuven.be*

***Received:*** *21 October 2021*
***Accepted:*** *18 January 2022*
***Published:*** *20 May 2022*

***Citation:***
*Naesens M, Budde K, Hilbrands L, Oberbauer R, Bellini MI, Glotz D, Grinyó J, Heemann U, Jochmans I, Pengel L, Reinders M, Schneeberger S and Loupy A (2022) Surrogate Endpoints for Late Kidney Transplantation Failure. Transpl Int 35:10136. doi: 10.3389/ti.2022.10136*

## INTRODUCTION

Key primary endpoints in kidney transplantation are recipient death, graft failure, biopsy-confirmed acute rejection, and graft (dys)function. These endpoints have clear roles in research that aims to improve short-term clinical outcomes after transplantation, and they are also the efficacy endpoints used most often in clinical trials (1). However, as improvement in short-term graft survival (by

**TABLE 1 |** Criteria for a valid surrogate endpoint, applied to potential surrogate endpoints in kidney transplantation.

| Criterion | Proteinuria | DSA | eGFR + proteinuria combined | Chronic graft histology | iBox score |
|---|---|---|---|---|---|
| Disease process (graft failure) sufficiently understood | Confirmed | Confirmed | Confirmed | Confirmed | Confirmed |
| Biologic plausibility | Confirmed | Confirmed | Confirmed | Confirmed | Confirmed |
| Strength of consistency supporting relationship between surrogate marker and outcome | Confirmed | Confirmed | Not confirmed | Not confirmed | Confirmed |
| Treatment effects on surrogate endpoint predict treatment effects on clinical outcome of interest | Not confirmed | Not confirmed | Not confirmed | Not confirmed | Not confirmed |

*DSA, donor-specific antibody; eGFR, estimated glomerular filtration rate.*

inhibition of early rejection) does not translate into an equally long-term improvement in graft survival, and as graft failure is rare in the early years following transplantation, better predictors of long-term graft outcome are needed for use in randomized controlled trials (RCT).

If conditional marketing authorization could be considered for therapies that aim to improve long-term outcomes [see Naesens et al., this issue (2)], then the surrogate endpoint for graft failure (i.e., loss of graft function; excluding death with a functioning graft) for use in RCT needs clearer definition. A good surrogate endpoint should fulfill four criteria: 1) The disease process is sufficiently understood; 2) The surrogate endpoint has biologic plausibility; 3) The strength of the consistency supports the relationship between the surrogate marker and outcome; 4) Treatment effects on the surrogate endpoint predict treatment effects on the clinical outcome of interest (**Table 1**). In addition, the acceptability of a surrogate endpoint for conditional marketing authorization of new therapies also depends on a benefit–risk evaluation and/or public health aspects, such as a serious life-threatening disease with no other therapeutic option, difficulties with studying the (rare or delayed) clinical endpoint, and the availability of a large safety database (2).

## DEFINITION AND CAUSES OF GRAFT FAILURE

Graft failure/loss of graft function is defined as return to dialysis or pre-emptive re-transplantation. Death of the recipient with a functioning graft is typically a primary safety endpoint, but we do not recommend including this in a surrogate endpoint for kidney transplantation outcome because of the wide variety of underlying causes of death observed (e.g., malignancy, infection, cardiovascular disease), lack of relation to graft functional status, and very different risk factors compared with those for graft failure (3, 4). These causes of death are often influenced by immunosuppression (5).

Furthermore, death with a functioning graft is a competing risk to loss of graft function, as is also the case in chronic kidney disease (CKD). In CKD, censoring for death increasingly overestimated the risk of kidney failure over time from 7% at 5 years to 19% at 10 years, especially in people at heightened risk of death (6). Although it could be anticipated that this is also relevant in kidney transplantation, the impact of this competing risk on the accuracy of death-censored graft failure risk is poorly established.

Definitions of all-cause and overall graft failure are discussed elsewhere in this Special Issue (2); of note, in this document, "graft failure" denotes loss of graft function, not overall graft failure (which includes patient death as a reason for graft failure). Given that late graft failure (excluding death with a functioning graft) is often multifactorial (4), it is difficult to predict such failure accurately with a single marker; a composite marker may more fully reflect the heterogeneity. The most important causes of graft failure are acute or chronic T-cell mediated rejection (TCMR), antibody-mediated rejection (AMR), nonspecific chronic injury due to nephron loss (drug toxicity, metabolic and urological factors), calcineurin inhibitor toxicity, infection, and other medical events (cardiorenal problems, vascular disease, malignancy, postrenal causes) (7), as well as occurrence or recurrence of original kidney disease. Consequently, the following markers are associated with heightened risk of late graft functional decline and failure: measured glomerular filtration rate (GFR); estimated (e)GFR, slope of eGFR trajectory, and eGFR change; CKD stage; proteinuria; *de novo* (*dn*) donor-specific antibodies (DSA); AMR histology; interstitial fibrosis and tubular atrophy (IFTA); and transplant glomerulopathy (TG) (8, 9).

## SINGLE MARKERS AS SURROGATE ENDPOINT

Single surrogate markers of graft function may not fully reflect the complexity of graft failure and death in kidney transplantation because some background (donor or recipient) risk factors—such as age and pre-existing immunological risk, including pre-transplant DSA—also affect outcome and graft-function markers. Late graft failure is more complex than renal failure resultant from native kidney disease because of competing risks involved at different time points. For example, the ZEUS trial (phase III randomized trial of cyclosporine continuation vs. switch to everolimus at 4.5 months post-transplant) showed a slightly better GFR (the primary endpoint), but higher rates of DSA and AMR (with absence of effect on graft failure and increased risk of graft failure) in patients who developed *dn*DSA (7, 10). Furthermore, creating too stringent a

definition of factors such as change in eGFR would require studies with long duration and large patient populations, which are difficult to achieve (11).

Conversely, considering only minor changes in a surrogate endpoint, such as eGFR or transplant glomerulopathy, increases the risk of error. For example, in histological terms, new or worsening transplant glomerulopathy could be considered as a surrogate endpoint in clinical trials, but the intrinsic heterogeneity of this pathology and varied data on its association with death-censored graft survival (12) make it difficult to translate findings into predictions for late events. In addition, this parameter has neither been used, nor accepted, by health authorities.

Combining multiple factors into a well-validated model is therefore more likely to predict long-term outcome (and better reflect the complexity of late graft failure) than using single endpoints or combining few factors. Relatively short-term improvements in such a complex score ideally would translate into long-term improvements in graft survival. It is also important that a valid surrogate marker for a well-understood disease process should have biological plausibility and a consistent relationship with outcome. Finally, treatment effects that change the surrogate marker should also have impact on clinical outcome.

Here we review the putative surrogate endpoints, including composite endpoints for predicting long-term graft outcome (excluding death with a functioning graft), focusing on eGFR, proteinuria, histological lesions, DSA, and complex scoring systems (**Table 1**).

## GFR and eGFR

For in-depth discussion on the association between kidney function and graft failure, methodology for measuring kidney function and its validity as a primary endpoint for clinical trials, see Hilbrands et al. (13).

Because graft failure is intrinsically defined by functional parameters such as dialysis reinitiation or repeat transplantation, graft functional assessment is directly related to the true endpoint, graft failure. Any intervention that stabilizes long-term graft function will inherently decrease the incidence of graft failure, therefore graft function is a direct measure of graft failure.

Predicted graft survival based on 12-months eGFR correlates with observed graft survival (14); consequently, eGFR alone is potentially interesting as a surrogate marker for long-term graft failure. This parameter was applied in the only relatively recent organ transplantation study to show improved long-term outcome using a new treatment (belatacept) (15, 16). However, declining eGFR is a late and insensitive marker for late graft failure in heterogeneous populations (17). The initial injury processes contributing to late graft failure are subclinical, and not reflected by early decline in renal function. Consequently, the long-term predictive value of measures of early post-transplantation renal function is limited (17); such measures (including serum creatinine values and use of eGFR) are discussed elsewhere in this supplement (13).

Additional graft injuries may develop slowly over time: declining renal function is the ultimate consequence of nephron loss but does not capture causes of nephron injury. Also, compensatory hyperfiltration may obscure initial damage. Moreover, the static absolute level of eGFR is also related to donor (e.g., age, brain death, hypotension) and transplant (e.g., ischemia/reperfusion) factors that might reduce the number of functioning nephrons at transplantation; using a single eGFR measurement as a surrogate endpoint would not take these into account.

Clearly, GFR has limitations as a surrogate for late graft failure, since in the first year after transplantation it fails to capture ongoing disease processes that lead to late graft failure. Sensitive tools that better reflect the heterogeneity in causes of late graft failure are required.

## Proteinuria

In CKD research there is increasing interest in using degree of proteinuria as a surrogate endpoint: the proteinuria level directly relates to the underlying glomerular disease process, and strongly correlates with progression to end-stage renal disease (18, 19). Proteinuria is routinely measured after kidney transplantation (20, 21); severe proteinuria in the nephrotic range often reflects structural damage to the nephron and is therefore associated with graft outcome (8, 22, 23). Histological signs of structural abnormality are TG, microcirculatory inflammation, and *dn* or recurrent glomerular disease (8), all of which are important causes of late graft failure.

Post-transplantation proteinuria thus tends to indicate poor prognosis, independent of graft function as assessed by eGFR (8, 24, 25), but may also reflect disease processes beyond renal function. Similar to general-population studies, an analysis that prospectively adjudicated cardiovascular events showed that albuminuria was strongly associated not only with graft failure, but also with cardiovascular events and mortality (25). Proteinuria alone has not widely been included as a surrogate endpoint in interventional studies of kidney transplantation and correcting post-transplantation proteinuria has not been proven to reduce the rates of long-term graft failure in studies of antihypertensive medication use in transplant populations (26–30). Conversely, studies with mTOR inhibitors revealed increases in proteinuria that did not translate into increased rates of long-term graft failure (7, 31).

## Donor-Specific HLA Antibodies

Since the early days of clinical kidney transplantation, it has been recognized that antibodies directed against non-self human leukocyte antigen (HLA) could be extremely relevant for graft outcomes. A seminal study described the key features and potential impact of alloantibodies in transplantation, demonstrating that immediate catastrophic graft failure is more likely to happen in multiparous female patients or in people receiving second transplants, and is seen in up to 80% of cases where there was a “positive crossmatch” (i.e., reactivity of recipient serum against donor cells) (32). The researchers advocated that demonstration of preformed

**TABLE 2 |** Association between changes in DSA and graft outcome in kidney transplantation RCTs. No studies show that interventions that affect DSA predict long-term graft outcomes (55, 61–63).

| Study | Setting and intervention | Effect on DSA | Effect on graft outcome |
|---|---|---|---|
| Bray et al., 2018 (55) | Belatacept vs. cyclosporine in the BENEFIT and BENEFIT-EXT studies | Significantly lower risk of *dn*DSA development and lower MFI of these DSA | Significantly better overall graft failure but equal death-censored graft failure and AMR risk |
| Moreso et al., 2018 (61) | IVIG + rituximab for chronic AMR | No change in immunodominant DSA-MFI between baseline and 1 year | No change in renal function assessed by eGFR (underpowered study) |
| Eskandary et al., 2018 (62) | Bortezomib vs. placebo for treatment of late AMR | No change in DSA-MFI | No change in renal function assessed by eGFR or graft failure |
| Sautenet et al., 2016 (63) | Rituximab vs. placebo for AMR | Significantly decreased DSA-MFI | No effect of the intervention on graft function or graft survival (underpowered study) |

*AMR, antibody-mediated rejection; dn,* de novo*; DSA, donor-specific antibodies; eGFR, estimated glomerular filtration rate; IVIG, intravenous immunoglobulin; MFI, mean fluorescence intensity; RCT, randomized controlled trial.*

cytotoxic antibodies against the graft ("sensitization") contraindicates allocation of a proposed graft to the transplant candidate. This recommendation was rapidly adopted and, aside from patients successfully desensitized by empirical approaches, remains rigidly enforced, using complement-dependent cytotoxicity crossmatching (CDC-XM) (32, 33).

With time, however, it became clear that CDC-XM lacks sensitivity for detecting circulating DSA: not all clinically significant pre-transplant DSA are identified. This led to the development of sensitive solid-phase tests, such as Luminex® single-antigen bead (SAB) assays (R&D Systems Inc., Minneapolis, MN, United States), which detect low-level DSA when the CDC test is negative. The definition of HLA antibody specificity by SAB assays added complexity to transplant risk stratification, by revealing extensive heterogeneity in the pathogenic potential of HLA-DSA. It is now well established that patients with pretransplant DSA detected by SAB, even with a negative CDC crossmatch, are at substantial risk of AMR and graft failure (34–37). Flow cytometry cross-matching adds additional insight into the actual immunologic risk for such patients (38).

The role of circulating anti–HLA-DSA is increasingly recognized as a major contributing factor to AMR and long-term graft failure (39–41). However, the occurrence of newly formed *dn*DSA after transplantation further increases the risk of graft failure (42–47), and complement-fixing DSA are particularly associated with graft rejection and failure (48). Some immunosuppressants (e.g., belatacept) appear to inhibit the development of *dn*HLA-DSA (16), while others (e.g., mTOR inhibitors) can be associated with a higher frequency of *dn*HLA-DSA (49). Importantly, under-immunosuppression and patient nonadherence are important risk factors for *dn*HLA-DSA development (50).

The STAR working group, a collaboration between the American Society for Histocompatibility and Immunogenetics and the American Society of Transplantation (51), made recommendations on the definitions and utilization of HLA diagnostic testing. In Europe, the European Federation for Immunogenetics publishes standards for histocompatibility and immunogenetics testing (52). Limitations of Luminex SAB assays that have been described include their semiquantitative nature, the prozone effect, test variability, and the need for arbitrary cut-off values to determine positivity. There are also technical challenges; for example, thresholds for DSA positivity are poorly defined and inconsistent, with European immunogenetics groups proposing mean fluorescence intensity (MFI) cut-off values of >3,000 or >5,000 MFI (53) and US groups proposing 1,400 MFI, which requires validation (51). A consistent definition of such a cut-off value, to indicate presence or absence of HLA antibodies, is crucial if DSA is to be considered as a single endpoint in RCTs. In addition, SAB MFI should not be used as a quantitative assay since it has a relatively high coefficient of variation (51). Thus, current technology cannot determine antibody titers or the clinical and biological relevance of positive test results (51, 54). In addition, although pretransplant DSA and *dn*HLA antibody development are strongly associated with AMR and graft failure (43, 55–60), no studies show that interventions affecting DSA levels or specificities after transplantation predict long-term improvement in graft survival rates (**Table 2**) (54, 61–63).

*Post hoc* analyses of the BENEFIT and BENEFIT-EXT studies (phase III randomized trials of belatacept vs. cyclosporine) showed significant reductions in the risk of *dn*DSA occurrence (55) and best overall graft survival rates. However, numbers were too small to demonstrate that these effects were mediated through improved death-censored graft survival or decreased risk of AMR. In contrast, data from mTOR inhibitor conversion studies showed higher rates of DSA and AMR in groups treated with mTOR inhibitors, but during the observation period no overall effect on graft survival was noted (64, 65), although follow-up was short, and DSA status was often missing (65). Finally, although the RITUX ERAH RCT (randomized trial of rituximab vs. placebo in addition to plasma exchange, intravenous immunoglobulin and corticosteroids for the treatment of AMR) showed an effect of rituximab on DSA-MFI that did not translate into improved graft function or survival rate, this study was underpowered, so firm conclusions could not be made (66).

As identified in a systematic review (67), therapeutic strategies eliminating *dn*DSA, tested in RCTs that are sufficiently powered to assess long-term graft outcomes, are needed. Case series suggest that "impossible" transplants become possible with

**TABLE 3 |** HR (multivariate models) for graft failure according to graft histology, renal function, and proteinuria at time of biopsy, adjusted for donor age and time after transplantation (8,9).

| Parameter | | Adjusted HR (95% CI) | *p* value |
|---|---|---|---|
| Naesens et al., 2016 (N = 1,335 indication biopsies) (8) | | | |
| Proteinuria at time of biopsy | 0.3–1.0 vs. <0.3 g/24 h | 1.14 (0.81–1.60) | 0.50 |
| | 1.0–3.0 vs. <0.3 g/24 h | 2.17 (1.49–3.18) | <0.001 |
| | >3.0 vs. <0.3 g/24 h | 3.01 (1.75–5.18) | <0.001 |
| eGFR at time of biopsy | 30–45 vs. >45 ml/min/1.73 $m^2$ | 1.76 (0.59–5.30) | 0.31 |
| | 15–30 vs. >45 ml/min/1.73 $m^2$ | 5.53 (1.99–15.4) | 0.001 |
| | <15 vs. >45 ml/min/1.73 $m^2$ | 11.7 (4.17–33.0) | <0.001 |
| Microcirculation inflammation | g + ptc ≥2 vs. <2 | 1.36 (0.97–1.91) | 0.07 |
| IFTA grade | Banff grade 1 vs. 0 | 1.82 (1.25–2.64) | 0.002 |
| | Banff grade 2–3 vs. 0 | 3.45 (2.34–5.07) | <0.001 |
| Transplant glomerulopathy | Banff grade 1 vs. 0 | 1.00 (0.55–1.82) | 0.99 |
| | Banff grade 2–3 vs. 0 | 1.83 (1.11–3.04) | 0.02 |
| *De novo*/recurrent glomerular disease | Present vs. absent | 1.35 (0.84–2.19) | 0.22 |
| Polyomavirus-associated nephropathy | Present vs. absent | 5.51 (3.06–9.92) | <0.001 |
| Loupy et al., 2019 (N = 3,941 patients) (9) | | | |
| Time from transplant to evaluation (years) | | 1.08 (1.02–1.14) | 0.0051 |
| eGFR (mL/min/1.73 $m^2$) | | 0.96 (0.95–0.96) | <0.0001 |
| Proteinuria (log) | | 1.51 (1.40–1.63) | <0.0001 |
| IFTA | 0/1 | — | |
| | 2 | 1.14 (0.918–1.424) | |
| | 3 | 1.39 (1.083–1.773) | 0.0311 |
| Microcirculation inflammation (g + ptc) | 0–2 | — | |
| | 3–4 | 1.45 (1.121–1.876) | |
| | 5–6 | 1.83 (1.240–2.706) | 0.0010 |
| Interstitial inflammation and tubulitis (i + t) | 0–2 | — | |
| | ≥3 | 1.34 (1.061–1.684) | 0.0136 |
| Transplant glomerulopathy (cg) | 0 | | |
| | ≥1 | 1.47 (1.133–1.895) | 0.0036 |
| Anti–HLA-DSA MFI | <500 | — | |
| | ≥500 to 3,000 | 1.25 (0.965–1.606) | |
| | ≥3,000 to 6,000 | 1.72 (1.115–2.659) | |
| | ≥6,000 | 2.05 (1.472–2.860) | 0.0001 |

*cg, transplant glomerulopathy; CI, confidence interval; DSA, donor-specific antibodies; eGFR, estimated glomerular filtration rate; g, glomerulitis score; HLA, human leukocyte antigen; HR, hazard ratio; i, interstitial; IFTA, interstitial fibrosis and tubular atrophy; MFI, mean fluorescence intensity; ptc, peritubular capillaritis score; t, tubulitis score.*

pre-transplant desensitization of HLA antibodies (67), but this does not validate HLA-DSA levels or specificities as surrogates for long-term outcome.

In summary, only the development of *dn*HLA-DSA with a clear MFI signal could be a meaningful surrogate endpoint that is strongly associated with adverse outcomes such as AMR and graft failure. While *dn*DSA development is clearly associated with immunosuppression, patient nonadherence (especially under-immunosuppression) may also play a role. The development of *dn*HLA-DSA has not been formally tested or validated as a surrogate endpoint for studies that aim to reduce graft failure because of AMR. In addition, as graft failure is heterogeneous and often no HLA-DSA are involved, *dn*DSA occurrence is insufficient as a surrogate for late graft failure by causes other than AMR.

## COMBINED FUNCTIONAL MARKERS

The risk of adverse outcomes at a given eGFR certainly increases with higher levels of albuminuria. In addition, integrating proteinuria and eGFR assessment is a good predictor of graft outcome (24, 25); studies also demonstrate an independent association between graft outcome and eGFR or proteinuria (8, 68).

Although potentially interesting as surrogate marker, the performance of a model that integrates proteinuria and eGFR has not been further validated in transplantation (25). However, whether the combination of eGFR and proteinuria could be considered as a primary (rather than surrogate) endpoint in kidney transplantation, as it is in CKD, warrants further discussion. Indeed, in CKD, the KDIGO guideline on prognostication based on integration of eGFR and albuminuria is an accepted surrogate for outcome in clinical trials, but the European Medicines Agency (EMA)'s CHMP guideline for primary prevention (69) proposed two primary efficacy endpoints: prevention or slowing of decline in the level of renal function (defined as either time to occurrence of CKD 3 or incidence rate of CKD ≥3); and clinically meaningful and stable difference in GFR failure rate with or without prevention of proteinuria/albuminuria. A similar primary endpoint could be considered in kidney transplantation, and the US Food and Drug Administration already follows this approach (70). However, no RCT has been undertaken to demonstrate that changes in such a composite functional endpoint predict changes in long-term graft survival rates.

**TABLE 4 |** Value of composite scores as surrogacy for long-term graft survival (9, 14, 72–77).

| Study | Kasiske et al., 2010 (72) | Foucher et al., 2010 (73) | Moore et al., 2011 (74) | Schnitzler et al., 2012 (14) | Shabir et al., 2014 (75); Gonzales et al., 2016 (76) | Gonzales et al., 2016 (76) | Prémaud et al., 2017 (77) | Loupy et al., 2019 (9) |
|---|---|---|---|---|---|---|---|---|
| Parameter | USRDS Risk Prediction Tool | KTFS | LOTESS Composite Risk Score | USRDS Predictive Model | Birmingham Risk Score | Birmingham-Mayo Histology-Based Model | AdGFS | iBox Risk Prediction Score |
| Development set | USRDS registry data (N = 59,091) | Multicentre French registry (N = 2,169) | Multicentre national cohort study (N = 2,763) | USRDS registry data (N = 87,575) | Single-center UK data (N = 651) | Single-center US data (N = 1,465) | Single-center French data (N = 664) | French multicentre cohort (N = 4,000) |
| External validation | No | Yes (N = 317) | Yes (single UK center; N = 731) | No | Yes (2 European centers (N = 736, N = 787) and 1 Canadian center (N = 475); 1 US center N = 1,465) | No | Yes (2 other French centers; N = 896) | Yes; N = 3,557 (2,129 patients in 3 European centers; 1,428 in 3 North American centers) |
| Prediction time point | 12 months post-transplant | 12 months post-transplant | Variable time after 12 months post-transplant | 12 months post-transplant | 12 months post-transplant | 12 months post-transplant | Time adjusted (only for 'rejection') | Time adjusted |
| Outcome parameter | Overall graft failure at 5 years after transplantation | Death-censored graft failure at 8 years | Overall graft failure and death-censored graft failure over time; follow-up time not specified | Overall graft failure beyond 1 year post-transplant, up to 9 years | Overall graft failure and death-censored graft failure at 5 years post-transplant | Overall graft failure and death-censored graft failure at 5 years post-transplant | Death-censored graft failure beyond 2 years post-transplant, up to 10 years | Death-censored graft failure over time post-transplant, up to 7 years |
| Pre-transplant factors included in the model | Recipient age<br>Recipient race<br>Insurance<br>Cause of ESRD | Recipient sex<br>Recipient age<br># Previous transplantations<br>Donor creatinine | Recipient age<br>Recipient sex<br>Recipient race | A large array of donor and recipient demographic factors (N = 20) | Recipient age<br>Recipient sex<br>Recipient race | Recipient age<br>Recipient sex<br>Recipient race | Donor age<br>Pre-transplant non-DSA HLA antibodies | Yes, adjusted for all relevant factors |
| Post-transplant factors included in the model | eGFR at 12 months<br>Hospitalization | Serum creatinine<br>Acute rejection<br>Creatinine at 3 months 24-h proteinuria | eGFR at 12 months<br>eGFR evolution<br>Acute rejection<br>Serum urea at 12 months<br>Serum albumin | eGFR at 12 months<br>Acute rejection within first year | Acute rejection<br>eGFR<br>Serum albumin<br>UACR | Acute rejection<br>eGFR<br>UACR<br>Black ethnicity<br>Glomerulitis score<br>Tubular atrophy score | Serum creatinine<br>Proteinuria<br>*dn*DSA<br>Serum creatinine trajectory<br>Acute rejection | Time post-transplant eGFR<br>Proteinuria<br>Histology (IFTA, microcirculation inflammation, TG)<br>DSA-MFI |
| Prognostic accuracy | C-statistic 0.65–0.78 | ROC AUC 0.78 (0.73–0.80) | C-statistic 0.83 for death-censored graft failure; 0.70 for overall graft failure | Not reported | C-statistic 0.78–0.90 for death-censored failure; 0.75–0.81 for overall graft failure | C-statistic 0.90 for death-censored failure; 0.81 for overall graft failure | C-statistic at 10 years post-transplant 0.83 (0.76–0.89) | C-statistic 0.81 in development cohort, 0.81 in European validation cohort, 0.80 in US validation cohort |

(Continued on following page)

**TABLE 4 |** (*Continued*) Value of composite scores as surrogacy for long-term graft survival (9, 14, 72–77).

| Study | Kasiske et al., 2010 (72) | Foucher et al., 2010 (73) | Moore et al., 2011 (74) | Schnitzler et al., 2012 (14) | Shabir et al., 2014 (75); Gonzales et al., 2016 (76) | Gonzales et al., 2016 (76) | Prémaud et al., 2017 (77) | Loupy et al., 2019 (9) |
|---|---|---|---|---|---|---|---|---|
| Calibration | Good | Not assessed | Good | Good | Good | Good | Good | Good |
| Limitations | No external validation set No data on DSA No data on proteinuria Prognostic accuracy moderate | Small validation set Validity not tested in other countries No data on DSA No data on rejection phenotype Limited prognostic accuracy | Small validation set Validity not tested in other countries No data on DSA No data on rejection phenotype Prediction time point variable | No external validation set No data on DSA No data on proteinuria No data on rejection phenotype | No data on rejection phenotype No data on DSA | No external validation set Data on DSA did not improve the model | Small validation sets and validity in other countries not tested Not tested in living donors or patients with pre-transplant DSA | Not yet prospectively implemented in an RCT |
| Tested in randomized trial data? | No | No | No | Yes, but calibration and validity as surrogacy for improved outcome by the intervention was not tested | No | No | No | Yes; validation in 3 RCTs; association with improved outcome not confirmed given lack of efficacy of the intervention |

*AdGFS, adjustable score for prediction of graft failure; dn,* de novo*; DSA, donor-specific antibodies; eGFR, estimated glomerular filtration rate; ESRD, end-stage renal disease; IFTA, interstitial fibrosis and tubular atrophy; KTFS, kidney transplantation failure score; LOTESS, long-term efficacy and safety surveillance; MFI, mean fluorescence intensity; RCT, randomized controlled trial; TG, transplant glomerulopathy; UACR, urine albumin to creatinine ratio; USRDS, United States Renal Data System.*

## COMPOSITE SCORES

Late graft failure (excluding death with a functioning graft) is a highly multifactorial state (4) that relates not only to early graft function, but also to subclinical injury processes including progressive IFTA or TG, drug toxicity, infections, medical events, recurrent disease, microvascular injury, and circulating DSA. Graft function is also highly dependent on pre-transplant donor/recipient risk factors (e.g., age, sex, delayed graft function), which further complicate the value of interpreting a single measurement of function as a surrogate for long-term outcome: studies show independent associations between these factors and graft failure in multivariate models (**Table 3**) (8, 9).

A systematic review evaluated models developed to predict graft failure in kidney transplantation recipients (71). Fourteen studies used predictors that were measured after transplantation; few studies integrated graft functional data such as proteinuria (n = 5) or serum creatinine/eGFR (n = 12), and none evaluated histology as part of the composite prediction model. Nineteen studies reported on the validity of the model in external datasets, several of which warrant in-depth assessment of their potential usefulness as surrogate endpoints for long-term graft failure excluding death with a functioning graft (14, 72–77); key features of these publications are listed in **Table 4**. Another study suggested a composite method for predicting graft failure; but because it included recipient death, it is less appropriate than other approaches as a potential surrogate endpoint for death-censored graft failure (78, 79).

In the study by Kasiske et al. (72), eGFR at 1 year was the only functional value included in the final model for prediction of 5-years graft failure, along with baseline recipient criteria and hospitalization within the first year following transplantation. However, this analysis was performed on a large registry (USRDS) that lacked crucial information on several clinical parameters. Furthermore, although the model showed good calibration, no independent validation was performed, and the impact of therapeutic interventions that aimed to reduce long-term graft failure was not tested. Moore et al. (74) restricted post-transplant factors in the model to eGFR and eGFR evolution, but nevertheless reached adequate discrimination and calibration for death-censored graft failure. External validation was restricted to a single center, and again the impact of therapeutic interventions was not evaluated. Importantly, the risk scores derived and tested in this study offered no prognostic superiority over basic metrics, such as eGFR or recipient age in isolation (74).

Foucher et al. proposed a clinical scoring system, built on the French DIVAT registry (3). The score was constructed at 1 year post transplantation, for prediction of graft failure at 8 years, and reached a C-statistic of 0.78. External validation was performed, but in a small dataset (n = 317). Other limitations included limited exportability, restriction to French transplant centers, and no inclusion of data on DSA and rejection subtypes or histological lesions. In addition, this score was built on observations at only

one time point. The potential of this prognostic score to be used as surrogacy for long-term graft failure was not tested in any RCT aiming to improve long-term outcome.

The first study to implement a previously developed risk score, in the context of a RCT aiming to improve long-term graft outcome, analyzed data from the USRDS registry (1995–2004) (14). Prediction models for all-cause graft survival were applied to participants in the BENEFIT and BENEFIT-EXT studies (phase III randomized trials of belatacept vs. cyclosporine), to determine whether the model could be used as a surrogate endpoint for late graft failure. Predicted and observed all-cause graft failures were well calibrated in standard- and expanded-criteria donor kidneys, as evaluated in the development cohort. Although data on model accuracy were lacking, aspects including eGFR and donor/recipient characteristics revealed a striking concordance between predicted and observed graft survival rates, when evaluated for 1-year outcome (14). However, when predicted survival estimates for 7 years post transplantation were compared with actual outcomes (16, 80), the predicted versus observed overall graft survival for the less-intensive group was 73.9 vs. 87.2%, and for the cyclosporine group was 69.0 vs. 78.3%. This illustrates that the calibration of the model for predicting longer-term survival was perhaps less than anticipated, which might be explained by the model being built on data obtained in an older era. As the surrogacy of the model established at 1 year for long-term graft failure was not directly confirmed, it is questionable whether it provides sufficient accuracy and calibration for use as a complex surrogate endpoint in future RCTs (14).

Shabir et al. developed a prediction model for 5-years graft failure using data from a single UK center, at 12 months post transplantation (75). The resultant risk scores were evaluated for prognostic utility (discrimination, calibration, and risk reclassification) in three independent cohorts in Europe and Canada. Recipient age, sex, and race; acute rejection rate; eGFR; serum albumin level; and urine albumin/creatinine ratio were included in scores for death-censored and overall graft failure. The rejection subtype was not further specified. In the validation cohorts, these scores showed good-to-excellent discrimination for death-censored transplant failure and moderate-to-good discrimination for overall transplant failure. Both scores demonstrated good calibration. Compared with eGFR in isolation, application of the scores resulted in statistically significant and clinically relevant risk reclassification for death-censored transplant failure [net reclassification improvement (NRI) 36.1–83.0%; all $p < 0.001$] and overall transplant failure (NRI 38.7–53.5%; all $p < 0.001$). Compared with the USRDS-based calculator, significant and relevant risk reclassification for overall transplant failure was seen (NRI 30.0%; $p < 0.001$) (75).

These scores have been externally validated (76): the risk model integrated 1-year histological and antibody data for prediction of graft failure at 5 years post transplantation in a single-center study (n = 1,465). The Birmingham Risk Score performed well, with good discrimination for recipients with or without graft failure 5 years after transplantation for both overall and death-censored graft failure (C-statistic 0.78 and 0.84, respectively), although this score has not been evaluated in an RCT designed to assess improvement of long-term graft outcome. Adding glomerulitis and interstitial fibrosis data to the Birmingham Risk Score improved the C-statistic for death-censored graft failure from 0.84 to 0.90, with further improved calibration and significant reclassification.

Decision-curve analyses aimed to determine how risk prediction could be improved when histological data were added to the clinical risk model proposed by Shabir et al. (75). However, this expanded model has not been independently validated and the impact of therapeutic interventions has not been evaluated. Prémaud et al. proposed a composite adjustable score for prediction of graft failure (AdGFS) using a conditional survival-tree analysis, undertaken using variables from patients transplanted between 1984 and 2011 in a French center (77). The analysis was based on serum creatinine and proteinuria at 12 months, *dn*DSA, serum creatinine cluster (creatinine value trajectories within the first year), acute rejection, donor age, and pre-transplant non-donor-specific HLA antibodies. Predictive performance of the AdGFS was good and the accuracy of the score at predicting graft failure remained high in the validation dataset, and in the external dataset (consisting of 896 patients from two other French centers, transplanted between 2002 and 2010). However, the study had limitations: the cohort did not represent current practice, there was no evaluation of the AdGFS response to therapies that aim to improve long-term graft outcome, validity in living donor kidney transplants and in recipients with pretransplant DSA was not tested, and data on DSA were lacking. In addition, international validation has not been performed.

## iBox

Loupy et al. developed the largest and only specifically designed multivariate model that predicts long-term death-censored graft failure: the iBox model was created after a study was undertaken in which parameters were collected from day of transplantation, to provide a holistic appraisal of potential risk factors (9). Their data showed that, among 7,557 kidney transplant recipients, 1,067 grafts failed (14.12%) in a median post-transplant follow-up of 7.12 years [interquartile range (IQR) 3.51–8.77] (9). In the derivation cohort, eight functional, histological, and immunological prognostic factors were found to be independently associated with death-censored graft failure. These were then combined into a risk prediction score that included the following parameters, in order of importance: eGFR; proteinuria:creatinine ratio; structural markers [Banff IFTA grade, microcirculation inflammation (Banff g + ptc), TG (Banff cg score), interstitial inflammation, and tubulitis (Banff i + t)]; MFI of the immunodominant HLA-DSA, and time from transplant to risk evaluation. The risk prediction score exhibited accurate calibration and discrimination (0.81 derivation and 0.80–0.81 in validation cohorts) (9). The performance of this multivariate model was validated in cohorts from three European and three North American centers (9). Importantly, testing the iBox model involved unselected patient cohorts, covering all potential clinical scenarios.

The iBox model was accurate when assessed independently of time since transplant, was validated in different clinical scenarios, and outperformed a risk score based solely on eGFR, proteinuria and HLA-DSA, not including histological lesions (**Table 5**). The risk prediction score was also slightly superior to the conventional graft monitoring model based on eGFR and proteinuria

**TABLE 5 |** Risk prediction score performance for iBox when assessed in different clinical scenarios and subpopulations (9).

| Risk score performance assessment | Risk model performance (C-statistic) | 95% bootstrap percentile CI |
|---|---|---|
| Functional and immunological parameters (without histology) | 0.79 | (0.77–0.81) |
| Histology diagnoses instead of Banff lesions grading | 0.76 | (0.74–0.81) |
| Stable patients (protocol biopsy) | 0.81 | (0.77–0.86) |
| Unstable patients (indication biopsy) | 0.80 | (0.78–0.82) |
| First year post-transplant | 0.77 | (0.72–0.81) |
| After 1 year post-transplant | 0.84 | (0.82–0.87) |
| Living donors | 0.82 | (0.75–0.88) |
| Deceased donors | 0.80 | (0.78–0.82) |
| Highly sensitized recipients | 0.80 | (0.76–0.84) |
| Non–highly sensitized recipients | 0.81 | (0.79–0.83) |
| Adding transplant baseline characteristics‡ | 0.81 | (0.79–0.83) |
| Patients with anti-IL-2 receptor induction | 0.79 | (0.76–0.82) |
| Patients with antithymocyte globulin induction | 0.83 | (0.80–0.85) |
| African American population | 0.80 | (0.74–0.85) |
| Non-African American population | 0.84 | (0.80–0.89) |
| Recipient blood pressure profile post-transplant | 0.80 | (0.78–0.82) |
| Calcineurin inhibitor blood level at time of evaluation | 0.81 | (0.78–0.83) |

*CI, confidence interval; eGFR, estimated glomerular filtration rate; IL, interleukin.*

**TABLE 6 |** Clinical trials depicting population characteristics, clinical scenarios and interventions, and prognostic performance of the iBox risk score (62–64).

| Study | Trial ID | Design | Clinical scenario | Target population | n | Time post-transplant (y) of risk score evaluation median, IQR | Follow-up time post-transplant (y) median, IQR | Risk score C-stat |
|---|---|---|---|---|---|---|---|---|
| CERTITEM (64) | NCT 01079143 | Prospective, randomized, open-label, multicentre trial | Immuno-suppressive drug minimization | Recipients of renal transplants from a living or deceased donor | 194 | 0.94 0.92–0.98 | 6.62 2.82–7.34 | 0.88 |
| RITUX ERAH (63) | EudraCT 2007-003213-13 | Prospective, randomized, multicentre, double-blind, placebo-controlled trial | AMR treatment (pre-existing DSA) | Recipients of renal transplants from a living or deceased donor with diagnosis of aAMR | 38 | 0.74 0.53–1.10 | 6.63 4.03–7.69 | 0.77 |
| BORTEJECT (62) | NCT 01873157 | Prospective, randomized, placebo-controlled, double-blind, single-center trial | AMR treatment (*dn*DSA) | Recipients of renal transplants from a living or deceased donor with post-transplant *dn*DSA detection | 44 | 6.61 4.04–15.41 | 7.75 5.32–16.41 | 0.94 |

*A, acute/active; AMR, antibody-mediated rejection; dn,* de novo*; DSA, donor-specific antibodies; IQR, interquartile range.*

assessments in terms of prediction capability; this was further demonstrated by a continuous NRI of 0.228 for the multivariate model compared with the functional model (95% confidence interval 0.174–0.290; $p < 0.0001$). In less-informed datasets, the new algorithm still performed with high accuracy (**Table 5**) (9).

The accuracy of the iBox risk score to predict long-term graft failure (9) was confirmed in *post hoc* analyses of data from three RCTs (**Table 6**) (62–64). Interventions performed in these studies affected the risk score, indicating that iBox adjusts to treatment effects. As the three RCTs did not significantly improve long-term graft outcome in the intervention group, the surrogacy of improvement of the score for predicting improvement of long-term graft survival could not be established directly. However, in the calcineurin inhibitor-free study arm of the CERTITEM study (randomized trial of switch to everolimus vs. cyclosporine continuation) there was a significantly increased risk of developing *dn*DSA in the everolimus group, higher rates of clinical or subclinical rejection, and worse eGFR, all of which were associated with a numerically higher risk of graft failure (5.2 vs. 1.0%). This difference in graft failure failed to reach statistical significance because of low event rates and thus lack of power (64). Post-hoc analysis of the TRANSFORM study (randomized trial of everolimus with reduced exposure calcineurin inhibitor vs. standard-exposure calcineurin inhibitor with mycophenolic acid) (81) indicated that an adapted iBox model (not all parameters were available) confirmed the noninferiority of everolimus with reduced cyclosporine vs. mycophenolic acid with standard cyclosporine for immunosuppression (82). The model projected kidney allograft survival up to 11 years postrandomization. The potential suitability of the iBox risk

score as being a surrogate endpoint is further indicated by its general validity, good calibration in RCTs, adjustability over time (and in response to treatment), and its integration of risk factors that are well confirmed in the pathophysiology of (or trajectory toward) graft failure. The evolution after kidney transplantation should be considered as a multidimensional pathophysiology, which could not be identified by looking at one parameter at a time. Importantly, extensive validation through modeling different post-transplant treatment interventions appears to confirm the association between each component of the score and long-term graft failure. For example, the iBox takes account of how a drug might affect kidney function by interfering with renal haemodynamics and eGFR but reducing DSA occurrence. In the context of a clinical trial or immediate therapeutic intervention, each parameter in iBox is individually ranked in terms of the performance, discrimination, and calibration of the risk score.

Statistical methodology used in iBox was directly derived from hazard ratio in the Cox analysis; other analyses (e.g., forms of machine learning) were tested but none of the models outperformed Cox, which is widely used in clinical research. The US Food and Drug Administration (FDA) has acknowledged the iBox as a "reasonably likely surrogate endpoint" biomarker to predict 5-years risk of graft failure in kidney transplantation (83). The developers are conducting further modeling to provide additional dimensions, including options for surrogacy, evaluation of its use as an early endpoint in clinical trials, and evaluation of its prognostic ability in subgroup analyses. The developers also plan to make the iBox an open-source platform and are preparing for the 507 drug-development tool qualification process, GDPR compliance, and other aspects of cybersecurity.

Several limitations of the iBox risk score should be noted. Firstly, the method is only useful for prediction of death-censored graft failure: adding death with a functioning graft as a safety endpoint remains necessary. The decision to use the iBox score for predicting death-censored graft failure rather than overall graft failure (including death with a functioning graft) was made because recipient death and loss of graft function have very different causes (3, 4, 71, 84). All-cause graft failure is usually multifactorial and needs a specific design with transplant characteristics, donor characteristics, and factors related to recipient's comorbidities at time of transplant and thereafter. In sensitivity analyses of the iBox study using competing risk regression models, allograft survival analyses performed in the final iBox model were not affected by competition with patient death.

Next, although the accuracy of the iBox model was maintained irrespective of whether histology was included as individual Banff lesion grades or histology diagnoses, scoring of individual histological lesions included in the composite score is hampered by reproducibility issues and interobserver variability. This limitation is relevant for any scoring system that includes histological parameters, is not specific for the iBox risk score, and needs to be addressed and mitigated in individual clinical trial designs and logistics. In addition, although the iBox score remained accurate across different centers using different methods of tissue typing and HLA antibody profiling, including the MFI of DSA means that this method is impacted by concerns relating to the absolute value of DSA-MFI, which is a semiquantitative rather than quantitative test. This must also be carefully addressed in clinical trial design.

With current evidence, we believe that the approach of multivariate models could be proposed as a surrogate marker for (death-censored) graft failure, since it considers the heterogeneity of causes of graft failure (excluding patient death with a functioning graft). Although it has not yet been shown in randomized trials that improvements in surrogate score actually predict improvements in long-term graft survival, the iBox model is the best-performing and best-validated algorithm to date (**Table 6**).

## CONCLUSIONS

- It is difficult to promote single markers as surrogate endpoints for late graft failure:
  - GFR has limitations, since the early course of graft function fails to capture ongoing subclinical disease processes. More sensitive tools are required that reflect heterogeneity in causes of late graft failure.
  - Early proteinuria is associated with late graft failure but has not been proposed or tested as a surrogate endpoint in kidney transplantation.
  - Combining GFR and proteinuria has a better association with graft failure than either factor separately, but its potential validity as a surrogate endpoint has not been tested.
  - Development of *dn*HLA-DSA is associated with graft failure but has not been formally tested or validated as a surrogate endpoint in studies that aim to reduce graft failure caused by AMR. As graft failure also occurs in the absence of AMR, *dn*DSA occurrence is insufficient as a surrogate for late graft failure by causes other than AMR.
- AMR and TCMR are primary endpoints for kidney transplantation clinical trials, which diminishes the need to pursue their validation as surrogate endpoints for late graft failure.
- Death of the recipient with a functioning graft is typically a primary safety endpoint:
  - Death of the recipient with a functioning graft is a competing risk for graft failure, but the impact of this competing risk on the accuracy of predictive models is poorly described.
  - We recommend not to include recipient death in a surrogate endpoint for late graft failure because of the wide variety of underlying causes of a death observed, different to the causes of graft failure.
- Several composite scores have been proposed and could be useful surrogate endpoints for interventional studies evaluating late graft failure.

- The iBox model is already a well-validated composite score that illustrates the robustness of this integrative approach, although further evaluations are in progress.

## Scientific Advice From the Commitee for Medicinal Products (CHMP) of the European Medicines Agency (EMA) for Human Use Regarding These Conclusions

- The CHMP acknowledged that the ESOT proposes to combine several factors into a single well-validated model as a surrogate endpoint to predict long-term outcome. A surrogate would be expected to fulfill the following three criteria:
  - Show biological relationship to the clinical outcome.
  - Demonstrate, in epidemiological studies, prognostic value of the surrogate for clinical outcome.
  - Provide evidence from clinical trial settings that treatment effect on the surrogate corresponds to effect on clinical outcome.
- The CHMP noted that ESOT introduces the iBox model (9) to predict long-term kidney graft failure at 3, 5, and 7 years, based on the following factors:
  - Time from transplant to risk evaluation after transplantation.
  - eGFR.
  - Proteinuria.
  - Banff IFTA grade, g + ptc, cg, and i + t scores.
  - MFI of donor-specific HLA antibodies.
- Based on ESOT´s position and the publication by Loupy et al. (9), the score appears to be designed as a risk calculation score and validated as such in separate cohorts. As such, the iBox score could provide an important contribution to the stratification of participants of clinical trials of transplantation.
- It is not clear if the third criterion above has been fulfilled, i.e., that treatment effect measured via iBox translates into corresponding effect on clinical outcome, i.e., graft failure. Furthermore, the following issues need to be addressed:
  - The statistical model and iBox algorithm were not presented and the relative contribution of each factor of the model was not evident; several factors of the iBox are also interrelated, e.g., histological diagnosis and the various histological lesions.
  - "Time from transplant" is an important prognostic marker but is never affected by therapy, therefore it cannot predict the effect of therapy on clinical outcome.
  - Outcome of iBox included death-censored graft failure, which is not a robust and favored clinical endpoint to show surrogacy, as there are several limitations in using the score without additional sensitivity analyses.
- ESOT showed the correlation of each variable in the final iBox model to death-censored functional outcome, a density plot of iBox evaluations post transplantation and the hazard ratio of each factor of the model.
  - Sensitivity analysis of the iBox indicate that graft survival analysis was not affected by competition with patient death.
  - ESOT noted that all-cause graft failure was multifactorial, with very different risk factors than death-censored graft failure, where grafts from patients who died with a functioning graft, were defined as functional grafts in the model.
  - However, ESOT acknowledged the concern regarding the importance of all-cause mortality in clinical trials of kidney transplantation for regulatory purposes and proposed to include this as part of safety or composite endpoints.
  - ESOT outlined the plans to further explore these issues with the FDA, including the preparation for a Drug Development Tool (DDT) qualification process.
- For the time being iBox is not qualified as a surrogate endpoint for regulatory purposes and thus cannot be proposed *a priori* to be used in clinical practice to guide decision making.
  - Based on the high-level data provided, CHMP notes that there are still certain limitations in applying the iBox score for regulatory purposes: the applicability of this score seems limited to certain determinants of kidney graft and the death-censored functional aspect.
  - A formal EMA Qualification of Novel Methodologies procedure for the finalized iBox as a surrogate marker would be very relevant way forward and is recommended.

## AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication. This article is one of a series of papers developed from the Broad Scientific Advice request, submitted to the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) by the European Society for Organ Transplantation (ESOT) in 2020: interactions between the EMA and ESOT regarding this request began in 2016. For the present article, through virtual and face-to-face discussions, the working group on surrogate endpoints in kidney transplantation developed the ESOT position on the core question 'Does CHMP agree that long-term outcome after kidney transplantation is an area of unmet medical need, for which conditional marketing authorization procedures should be considered, to facilitate timely access to new therapies? If so, does CHMP agree with the proposed surrogate endpoints for clinical trials for therapies requiring conditional marketing authorization?' The Centre for Evidence in Transplantation provided support with specific data extraction requests: these literature searches formed the basis of evidence used in the Broad Scientific Advice request and the present article. Input into the working group's output was provided from all ESOT members involved in the Broad Scientific Advice request process. The present article was adapted by MN from the final Broad Scientific Advice request submission (June 2020), presentation documents and minutes of the meeting between ESOT and the

CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020). All drafts of the article were circulated to all co-authors for review and approval before submission for publication.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

KB has received honoraria and/or research funding from Alexion, Astellas, Bristol Myers Squibb, Chiesi, Fresenius, Hansa, Hexal, Merck, Novartis, Otsuka, Pfizer, Roche, Sandoz, Siemens, and Veloxis. LH reports speaker fees from Astellas, consultancy and research support from Chiesi, consultancy for Novartis, and research support from Sandoz. RO has received grants/research support from Amgen, Astellas, and Chiesi; and speakers' bureaux/honoraria from Amgen, Astellas, Chiesi, Hansa, Neovii, Novartis, and Teva. DG reports consultancy agreements and scientific boards for AstraZeneca, BMS, Hansa, and Sanofi. UH has received grants/research support from Baxter, Chiesi, and Neovii; speakers' bureaux/honoraria from Chiesi and Hansa; and consulting fees from Astellas, Hansa, Neovii, Novartis, and Teva. IJ's institution has received speaker's fees from XVIVO Perfusion. MR has received lecture fees from Astellas and Chiesi; and research grant support (paid to institution) from Astellas and Chiesi for investigator-initiated studies. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. Sautenet B, TongChapman AJR, Chapman JR, Warrens AN, Rosenbloom D, Wong G, et al. Range and Consistency of Outcomes Reported in Randomized Trials Conducted in Kidney Transplant Recipients: a Systematic Review. *Transplantation* (2018) 102:2065–71. doi:10.1097/tp.0000000000002278
2. Naesens M, Loupy A, Hilbrands L, Oberbauer R, Bellini MI, Glotz D, et al. Rationale for Surrogate Endpoints and Conditional Marketing Authorization of New Therapies for Kidney Transplantation. *Transpl Int* (2022). 35:10137. doi:10.3389/ti.2022.10137
3. Van Loon E, Senev A, Lerut E, Coemans M, Callemeyn J, Van Keer JM, et al. Assessing the Complex Causes of Kidney Allograft Loss. *Transplantation* (2020) 104:2557–66. doi:10.1097/tp.0000000000003192
4. Mayrdorfer M, Liefeldt L, Wu K, Rudolph B, Zhang Q, Friedersdorff F, et al. Exploring the Complexity of Death-Censored Kidney Allograft Failure. *J Am Soc Nephrol* (2021) 32:1513–26. doi:10.1681/asn.2020081215
5. Opelz G, Döhler B. Association between Steroid Dosage and Death with a Functioning Graft after Kidney Transplantation. *Am J Transplant* (2013) 13: 2096–105. doi:10.1111/ajt.12313
6. Ravani P, Fiocco M, Liu P, Quinn RR, Hemmelgarn B, James M, et al. Influence of Mortality on Estimating the Risk of Kidney Failure in People with Stage 4 CKD. *J Am Soc Nephrol* (2019) 30:2219–27. doi:10.1681/asn.2019060640
7. Budde K, Lehner F, Sommerer C, Reinke P, Arns W, Eisenberger U, et al. Five-year Outcomes in Kidney Transplant Patients Converted from Cyclosporine to Everolimus: the Randomized ZEUS Study. *Am J Transplant* (2014) 15:119–28. doi:10.1111/ajt.12952
8. Naesens M, Lerut E, Emonds M-P, Herelixka A, Evenepoel P, Claes K, et al. Proteinuria as a Noninvasive Marker for Renal Allograft Histology and Failure: an Observational Cohort Study. *J Am Soc Nephrol* (2016) 27:281–92. doi:10.1681/asn.2015010062
9. Loupy A, Aubert O, Orandi BJ, Naesens M, Bouatou Y, Raynaud M, et al. Prediction System for Risk of Allograft Loss in Patients Receiving Kidney Transplants: International Derivation and Validation Study. *BMJ* (2019) 366: l4923. doi:10.1136/bmj.l4923
10. O'Connell PJ, Kuypers DR, Mannon RB, Abecassis M, Chadban SJ, Gill JS, et al. Clinical Trials for Immunosuppression in Transplantation: the Case for Reform and Change in Direction. *Transplantation* (2017) 101:1527–34. doi:10.1097/TP.0000000000001648
11. Stegall MD, Morris RE, Alloway RR, Mannon RB. Developing New Immunosuppression for the Next Generation of Transplant Recipients: the Path Forward. *Am J Transplant* (2016) 16:1094–101. doi:10.1111/ajt.13582
12. Aubert O, Higgins S, Bouatou Y, Yoo D, Raynaud M, Viglietti D, et al. Archetype Analysis Identifies Distinct Profiles in Renal Transplant Recipients with Transplant Glomerulopathy Associated with Allograft Survival. *J Am Soc Nephrol* (2019) 30:625–39. doi:10.1681/asn.2018070777
13. Hilbrands L, Budde K, Bellini MI, Diekmann F, Furian L, Grinyo J, et al. Allograft Function as Endpoint for Clinical Trials in Kidney Transplantation. *Transpl Int* (2022) 35:10139. doi:10.3389/ti.2022.10139
14. Schnitzler MA, Lentine KL, Axelrod D, Gheorghian A, You M, Kalsekar A, et al. Use of 12-Month Renal Function and Baseline Clinical Factors to Predict Long-Term Graft Survival. *Transplantation* (2012) 93:172–81. doi:10.1097/tp.0b013e31823ec02a
15. Larsen CP, Grinyó J, Medina-Pestana J, Vanrenterghem Y, Vincenti F, Breshahan B, et al. Belatacept-based Regimens versus a Cyclosporine A-Based Regimen in Kidney Transplant Recipients: 2-year Results from the BENEFIT and BENEFIT-EXT Studies. *Transplantation* (2010) 90:1528–35. doi:10.1097/tp.0b013e3181ff87cd
16. Vincenti F, Rostaing L, Grinyo J, Rice K, Steinberg S, Gaite L, et al. Belatacept and Long-Term Outcomes in Kidney Transplantation. *N Engl J Med* (2016) 374:333–43. doi:10.1056/nejmoa1506027
17. Kaplan B, Schold and J, Meier-Kriesche H-U. Poor Predictive Value of Serum Creatinine for Renal Allograft Loss. *Am J Transplant* (2003) 3:1560–5. doi:10.1046/j.1600-6135.2003.00275.x
18. Cravedi P, Ruggenenti P, Remuzzi G. Proteinuria Should Be Used as a Surrogate in CKD. *Nat Rev Nephrol* (2012) 8:301–6. doi:10.1038/nrneph.2012.42

19. Avasare RS, Radhakrishnan J. Proteinuria as a Surrogate Marker for Renal Outcome: Are We There yet? *Kidney Int* (2015) 88:1228–30. doi:10.1038/ki.2015.301
20. Nankivell BJ, Kuypers DR. Diagnosis and Prevention of Chronic Kidney Allograft Loss. *Lancet* (2011) 378:1428–37. doi:10.1016/s0140-6736(11)60699-5
21. Legendre C, Anglicheau D. Proteinuria in Kidney Transplantation: an Ongoing story. *Nat Rev Nephrol* (2013) 9:251–2. doi:10.1038/nrneph.2013.61
22. Amer H, Fidler ME, Myslak M, Morales P, Kremers WK, Larson TS, et al. Proteinuria after Kidney Transplantation, Relationship to Allograft Histology and Survival. *Am J Transplant* (2007) 7:2748–56. doi:10.1111/j.1600-6143.2007.02006.x
23. Halimi J-M, Buchler M, Al-Najjar A, Laouad I, Chatelet V, Marlière J-F, et al. Urinary Albumin Excretion and the Risk of Graft Loss and Death in Proteinuric and Non-proteinuric Renal Transplant Recipients. *Am J Transplant* (2007) 7:618–25. doi:10.1111/j.1600-6143.2007.01665.x
24. Lam NN, Tonelli M, Lentine KL, Hemmelgarn B, Ye F, Wen K, et al. Albuminuria and Posttransplant Chronic Kidney Disease Stage Predict Transplant Outcomes. *Kidney Int* (2017) 92:470–8. doi:10.1016/j.kint.2017.01.028
25. Weiner DE, Park M, Tighiouart H, Joseph AA, Carpenter MA, Goyal N, et al. Albuminuria and Allograft Failure, Cardiovascular Disease Events, and All-Cause Death in Stable Kidney Transplant Recipients: a Cohort Analysis of the FAVORIT Trial. *Am J Kidney Dis* (2019) 73:51–61. doi:10.1053/j.ajkd.2018.05.015
26. Opelz G, Zeier M, Laux G, Morath C, Döhler B. No Improvement of Patient or Graft Survival in Transplant Recipients Treated with Angiotensin-Converting Enzyme Inhibitors or Angiotensin II Type 1 Receptor Blockers: a Collaborative Transplant Study Report. *J Am Soc Nephrol* (2006) 17:3257–62. doi:10.1681/asn.2006050543
27. Philipp T, Martinez F, Geiger H, Moulin B, Mourad G, Schmieder R, et al. Candesartan Improves Blood Pressure Control and Reduces Proteinuria in Renal Transplant Recipients: Results from SECRET. *Nephrol Dial Transplant* (2010) 25:967–76. doi:10.1093/ndt/gfp581
28. Mandelbrot DA, Alberú J, Barama A, Marder BA, Silva HT, Jr, Flechner SM, et al. Effect of Ramipril on Urinary Protein Excretion in Maintenance Renal Transplant Patients Converted to Sirolimus. *Am J Transplant* (2015) 15:3174–84. doi:10.1111/ajt.13384
29. Knoll GA, Fergusson D, Chassé M, Hebert P, Wells G, Tibbles LA, et al. Ramipril versus Placebo in Kidney Transplant Patients with Proteinuria: a Multicentre, Double-Blind, Randomised Controlled Trial. *Lancet Diabetes Endocrinol* (2016) 4:318–26. doi:10.1016/s2213-8587(15)00368-x
30. Hiremath S, Fergusson DA, Fergusson N, Bennett A, Knoll GA. Renin-angiotensin System Blockade and Long-Term Clinical Outcomes in Kidney Transplant Recipients: a Meta-Analysis of Randomized Controlled Trials. *Am J Kidney Dis* (2017) 69:78–86. doi:10.1053/j.ajkd.2016.08.018
31. Budde K, Becker T, Arns W, Sommerer C, Reinke P, Eisenberger U, et al. Everolimus-based, Calcineurin-inhibitor-free Regimen in Recipients of De-novo Kidney Transplants: an Open-Label, Randomised, Controlled Trial. *Lancet* (2011) 377:837–47. doi:10.1016/s0140-6736(10)62318-5
32. Patel R, Terasaki PI. Significance of the Positive Crossmatch Test in Kidney Transplantation. *N Engl J Med* (1969) 280:735–9. doi:10.1056/nejm196904032801401
33. Gebel HM, Bray RA, Nickerson P. Pre-transplant Assessment of Donor-Reactive, HLA-specific Antibodies in Renal Transplantation: Contraindication vs. Risk. *Am J Transplant* (2003) 3:1488–500. doi:10.1046/j.1600-6135.2003.00273.x
34. Ziemann M, Altermann W, Angert K, Arns W, Bachmann A, Bakchoul T, et al. Preformed Donor-specific HLA Antibodies in Living and Deceased Donor Transplantation. *Clin J Am Soc Nephrol* (2019) 14:1056–66. doi:10.2215/cjn.13401118
35. Michielsen LA, Wisse BW, Kamburova EG, Verhaar MC, Joosten I, Allebes WA, et al. A Paired Kidney Analysis on the Impact of Pre-transplant Anti-HLA Antibodies on Graft Survival. *Nephrol Dial Transplant* (2019) 34:1056–63. doi:10.1093/ndt/gfy316
36. Mohan S, Palanisamy A, Tsapepas D, Tanriover B, Crew RJ, Dube G, et al. Donor-specific Antibodies Adversely Affect Kidney Allograft Outcomes. *J Am Soc Nephrol* (2012) 23:2061–71. doi:10.1681/asn.2012070664
37. Zecher D, Bach C, Staudner C, Böger CA, Bergler T, Banas B, et al. Characteristics of Donor-specific Anti-HLA Antibodies and Outcome in Renal Transplant Patients Treated with a Standardized Induction Regimen. *Nephrol Dial Transplant* (2017) 32:730–7. doi:10.1093/ndt/gfw445
38. Konvalinka A, Tinckam K. Utility of HLA Antibody Testing in Kidney Transplantation. *J Am Soc Nephrol* (2015) 26:1489–502. doi:10.1681/asn.2014080837
39. Nankivell BJ, Alexander SI. Rejection of the Kidney Allograft. *N Engl J Med* (2010) 363:1451–62. doi:10.1056/nejmra0902927
40. Lefaucheur C, Loupy A, Hill GS, Andrade J, Nochy D, Antoine C, et al. Preexisting Donor-specific HLA Antibodies Predict Outcome in Kidney Transplantation. *J Am Soc Nephrol* (2010) 21:1398–406. doi:10.1681/asn.2009101065
41. Senev A, Coemans M, Lerut E, Van Sandt V, Daniëls L, Kuypers D, et al. Histological Picture of Antibody-mediated Rejection without Donor-specific anti-HLA Antibodies: Clinical Presentation and Implications for Outcome. *Am J Transplant* (2019) 19:763–80. doi:10.1111/ajt.15074
42. Wiebe C, Gibson IW, Blydt-Hansen TD, Karpinski M, Ho J, Storsley LJ, et al. Evolution and Clinical Pathologic Correlations of De Novo Donor-specific HLA Antibody post Kidney Transplant. *Am J Transplant* (2012) 12:1157–67. doi:10.1111/j.1600-6143.2012.04013.x
43. Aubert O, Loupy A, Hidalgo L, Duong van Huyen J-P, Higgins S, Viglietti D, et al. Antibody-Mediated Rejection Due to Preexisting versusDe NovoDonor-specific Antibodies in Kidney Allograft Recipients. *J Am Soc Nephrol* (2017) 28:1912–23. doi:10.1681/asn.2016070797
44. Cooper JE, Gralla J, Cagle L, Goldberg R, Chan L, Wiseman AC. Inferior Kidney Allograft Outcomes in Patients with De Novo Donor-specific Antibodies Are Due to Acute Rejection Episodes. *Transplantation* (2011) 91:1103–9. doi:10.1097/tp.0b013e3182139da1
45. Ginevri F, Nocera A, Comoli P, Innocente A, Cioni M, Parodi A, et al. PosttransplantDe NovoDonor-specific HLA Antibodies Identify Pediatric Kidney Recipients at Risk for Late Antibody-Mediated Rejection. *Am J Transplant* (2012) 12:3355–62. doi:10.1111/j.1600-6143.2012.04251.x
46. Zhang R. Donor-specific Antibodies in Kidney Transplant Recipients. *Clin J Am Soc Nephrol* (2018) 13:182–92. doi:10.2215/cjn.00700117
47. Mannon RB, Askar M, Jackson AM, Newell K, Mengel M. Meeting Report of the STAR-Sensitization in Transplantation Assessment of Risk: Naïve Abdominal Transplant Organ Subgroup Focus on Kidney Transplantation. *Am J Transplant* (2018) 18:2120–34. doi:10.1111/ajt.14977
48. Bouquegneau A, Loheac C, Aubert O, Bouatou Y, Viglietti D, Empana JP, et al. Complement-activating Donor-specific Anti-HLA Antibodies and Solid Organ Transplant Survival: A Systematic Review and Meta-Analysis. *PLoS* (2018) 15:e1002572. doi:10.1371/journal.pmed.1002572
49. Grimbert P, Thaunat O. mTOR Inhibitors and Risk of Chronic Antibody-Mediated Rejection after Kidney Transplantation: where Are We Now? *Transpl Int* (2017) 30:647–57. doi:10.1111/tri.12975
50. Gokoel SRM, Gombert-Handoko KB, Zwart TC, van der Boog PJM, Moes DJAR, de Fijter JW. Medication Non-adherence after Kidney Transplantation: A Critical Appraisal and Systematic Review. *Transplant Rev* (2020) 34:100511. doi:10.1016/j.trre.2019.100511
51. Tambur AR, Wiebe C. HLA Diagnostics. *Transplantation* (2018) 102:S23–S30. doi:10.1097/tp.0000000000001817
52. European Federation of Immunogenetics. EFI Standards 2019/2020 (2019). Available at: https://efi-web.org/committees/standards-committee (Accessed January 14, 2022).
53. Süsal C, Seidl C, Schönemann C, Heinemann FM, Kauke T, Gombos P, et al. Determination of Unacceptable HLA Antigen Mismatches in Kidney Transplant Recipients: Recommendations of the German Society for Immunogenetics. *Tissue Antigens* (2015) 86:317–23. doi:10.1111/tan.12682
54. Karahan GE, Claas FHJ, Heidt S. Technical Challenges and Clinical Relevance of Single Antigen Bead C1q/C3d Testing and IgG Subclass Analysis of Human Leukocyte Antigen Antibodies. *Transpl Int* (2018) 31:1189–97. doi:10.1111/tri.13327
55. Bray RA, Gebel HM, Townsend R, Roberts ME, Polinsky M, Yang L, et al. De Novo donor-specific Antibodies in Belatacept-Treated vs Cyclosporine-Treated Kidney-Transplant Recipients: Post Hoc Analyses of the Randomized Phase III BENEFIT and BENEFIT-EXT Studies. *Am J Transplant* (2018) 18:1783–9. doi:10.1111/ajt.14721

56. Wiebe C, Gibson IW, Blydt-Hansen TD, Pochinco D, Birk PE, Ho J, et al. Rates and Determinants of Progression to Graft Failure in Kidney Allograft Recipients WithDe NovoDonor-specific Antibody. *Am J Transplant* (2015) 15:2921–30. doi:10.1111/ajt.13347
57. Wiebe C, Rush DN, Nevins TE, Birk PE, Blydt-Hansen T, Gibson IW, et al. Class II Eplet Mismatch Modulates Tacrolimus Trough Levels Required to Prevent Donor-specific Antibody Development. *J Am Soc Nephrol* (2017) 28: 335–62. doi:10.1681/asn.2017030287
58. Lachmann N, Terasaki PI, Budde K, Liefeldt L, Kahl A, Reinke P, et al. Anti-human Leukocyte Antigen and Donor-specific Antibodies Detected by Luminex Posttransplant Serve as Biomarkers for Chronic Rejection of Renal Allografts. *Transplantation* (2009) 87:1505–13. doi:10.1097/tp.0b013e3181a44206
59. Lachmann N, Niemann M, Reinke P, Budde K, Schmidt D, Halleck F, et al. Donor-Recipient Matching Based on Predicted Indirectly Recognizable HLA Epitopes Independently Predicts the Incidence ofDe NovoDonor-specific HLA Antibodies Following Renal Transplantation. *Am J Transplant* (2017) 17: 3076–86. doi:10.1111/ajt.14393
60. Liefeldt L, Budde K. Risk Factors for Cardiovascular Disease in Renal Transplant Recipients and Strategies to Minimize Risk. *Transpl Int* (2010) 23:1191–204. doi:10.1111/j.1432-2277.2010.01159.x
61. Moreso F, Crespo M, Ruiz JC, Torres A, Gutierrez-Dalmau A, Osuna A, et al. Treatment of Chronic Antibody Mediated Rejection with Intravenous Immunoglobulins and Rituximab: A Multicenter, Prospective, Randomized, Double-Blind Clinical Trial. *Am J Transplant* (2018) 18:927–35. doi:10.1111/ajt.14520
62. Eskandary F, Regele H, Baumann L, Bond G, Kozakowski N, Wahrmann M, et al. A Randomized Trial of Bortezomib in Late Antibody-Mediated Kidney Transplant Rejection. *J Am Soc Nephrol* (2018) 29:591–605. doi:10.1681/asn.2017070818
63. Sautenet B, Blancho G, Büchler M, Morelon E, Toupance O, Barrou B, et al. One-year Results of the Effects of Rituximab on Acute Antibody-Mediated Rejection in Renal Transplantation. *Transplantation* (2016) 100:391–9. doi:10.1097/tp.0000000000000958
64. Rostaing L, Hertig A, Albano L, Anglicheau D, Durrbach A, Vuiblet V, et al. Fibrosis Progression According to Epithelial-Mesenchymal Transition Profile: a Randomized Trial of Everolimus versus CsA. *Am J Transplant* (2015) 15: 1303–12. doi:10.1111/ajt.13132
65. De Fijter JW, Holdaas H, Øyen O, Sanders J-S, Sundar S, Bemelman FJ, et al. Early Conversion from Calcineurin Inhibitor- to Everolimus-Based Therapy Following Kidney Transplantation: Results of the Randomized ELEVATE Trial. *Am J Transplant* (2017) 17:1853–67. doi:10.1111/ajt.14186
66. Bailly E, Ville S, Blancho G, Morelon E, Bamoulid J, Caillard S, et al. An Extension of the RITUX-ERAH Study, Multicenter Randomized Clinical Trial Comparing Rituximab to Placebo in Acute Antibody-mediated Rejection after Renal Transplantation. *Transpl Int* (2020) 33:786–95. doi:10.1111/tri.13613
67. Sharma A, Lewis JR, Lim WH, Palmer S, Strippoli G, Chapman JR, et al. Renal Transplant Outcomes and De Novo Donor-specific Anti-human Leukocyte Antigen Antibodies: a Systematic Review. *Nephrol Dial Transplant* (2018) 33: 1472–80. doi:10.1093/ndt/gfy077
68. Loupy A, Vernerey D, Tinel C, Aubert O, Duong van Huyen J-P, Rabant M, et al. Subclinical Rejection Phenotypes at 1 Year post-transplant and Outcome of Kidney Allografts. *J Am Soc Nephrol* (2015) 26:1721–31. doi:10.1681/asn.2014040399
69. European Medicines Agency/CHMP/500825/2016. Guideline on the Clinical Investigation of Medicinal Products to Prevent Development/slow Progression of Chronic Renal Insufficiency. Available at https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-prevent-development/slow-progression-chronic-renal-insufficiency_en.pdf (Accessed December 20, 2021).
70. Levey AS, Inker LA, Matsushita K, Greene T, Willis K, Lewis E, et al. GFR Decline as an End point for Clinical Trials in CKD: a Scientific Workshop Sponsored by the National Kidney Foundation and the US Food and Drug Administration. *Am J Kidney Dis* (2014) 64:821–35. doi:10.1053/j.ajkd.2014.07.030
71. Kaboré R, Haller MC, Harambat J, Heinze G, Leffondré K. Risk Prediction Models for Graft Failure in Kidney Transplantation: a Systematic Review. *Nephrol Dial Transplant* (2017) 32:ii68–ii76. doi:10.1093/ndt/gfw405
72. Kasiske BL, Israni AK, Snyder JJ, Skeans MA, Peng Y, Weinhandl ED. A Simple Tool to Predict Outcomes after Kidney Transplant. *Am J Kidney Dis* (2010) 56:947–60. doi:10.1053/j.ajkd.2010.06.020
73. Foucher Y, Daguin P, Akl A, Kessler M, Ladrière M, Legendre C, et al. A Clinical Scoring System Highly Predictive of Long-Term Kidney Graft Survival. *Kidney Int* (2010) 78:1288–94. doi:10.1038/ki.2010.232
74. Moore J, He X, Shabir S, Hanvesakul R, Benavente D, Cockwell P, et al. Development and Evaluation of a Composite Risk Score to Predict Kidney Transplant Failure. *Am J Kidney Dis* (2011) 57:744–51. doi:10.1053/j.ajkd.2010.12.017
75. Shabir S, Halimi J-M, Cherukuri A, Ball S, Ferro C, Lipkin G, et al. Predicting 5-year Risk of Kidney Transplant Failure: a Prediction Instrument Using Data Available at 1 Year Posttransplantation. *Am J Kidney Dis* (2014) 63:643–51. doi:10.1053/j.ajkd.2013.10.059
76. Gonzales MM, Bentall A, Kremers WK, Stegall MD, Borrows R. Predicting Individual Renal Allograft Outcomes Using Risk Models with 1-year Surveillance Biopsy and Alloantibody Data. *J Am Soc Nephrol* (2016) 27: 3165–74. doi:10.1681/asn.2015070811
77. Prémaud A, Filloux M, Gatault P, Thierry A, Büchler M, Munteanu E, et al. An Adjustable Predictive Score of Graft Survival in Kidney Transplant Patients and the Levels of Risk Linked to De Novo Donor-specific Anti-HLA Antibodies. *PLoS One* (2017) 12:e0180236. doi:10.1371/journal.pone.0180236
78. Srinivas TR, Taber DJ, Su Z, Zhang J, Mour G, Northrup D, et al. Big Data, Predictive Analytics, and Quality Improvement in Kidney Transplantation: a Proof of Concept. *Am J Transplant* (2017) 17:671–81. doi:10.1111/ajt.14099
79. DuBay DA, Su Z, Morinelli TA, Baliga P, Rohan V, Bian J, et al. Development and Future Deployment of a 5 Years Allograft Survival Model for Kidney Transplantation. *Nephrology* (2019) 24:855–62. doi:10.1111/nep.13488
80. Vincenti F, Blancho G, Durrbach A, Grannas G, Grinyó J, Meier-Kriesche H-U, et al. Ten-year Outcomes in a Randomized Phase II Study of Kidney Transplant Recipients Administered Belatacept 4-weekly or 8-weekly. *Am J Transplant* (2017) 17:3219–27. doi:10.1111/ajt.14452
81. Pascual J, Berger SP, Witzke O, Tedesco H, Mulgaonkar S, Qazi Y, et al. Everolimus with Reduced Calcineurin Inhibitor Exposure in Renal Transplantation. *J Am Soc Nephrol* (2018) 29:1979–91. doi:10.1681/asn.2018010009
82. Aubert O, Divard G, Pascual J, Oppenheimer F, Sommerer C, Citterio F, et al. Application of the iBox Prognostication System as a Surrogate Endpoint in the TRANSFORM Randomised Controlled Trial: Proof-Of-Concept Study. *BMJ Open* (2021) 11:e052138. doi:10.1136/bmjopen-2021-052138
83. Critical Path Institute Transplant Therapeutics Consortium. The Integrative Box (iBox) Scoring System as a Reasonably Likely Surrogate Endpoint for Five-Year Risk of Allograft Loss in Kidney Transplant Recipients for Use in Clinical Trials Evaluating the Safety and Efficacy of Novel Immunosuppressive Drug Therapies (2020). Available at https://www.fda.gov/media/139301/download (Accessed December 20, 2021).
84. El-Zoghby ZM, Stegall MD, Lager DJ, Kremers WK, Amer H, Gloor JM, et al. Identifying Specific Causes of Kidney Allograft Loss. *Am J Transplant* (2009) 9: 527–35. doi:10.1111/j.1600-6143.2008.02519.x

ESOT | Transplant International

CONSENSUS REPORT
published: 20 May 2022
doi: 10.3389/ti.2022.10134

# Patient-Reported Outcomes as Endpoints in Clinical Trials of Kidney Transplantation Interventions

*Allison Tong[1†], Rainer Oberbauer[2†], Maria Irene Bellini[3], Klemens Budde[4], Fergus J. Caskey[5], Fabienne Dobbels[6], Liset Pengel[7], Lionel Rostaing[8], Stefan Schneeberger[9] and Maarten Naesens[10*]*

[1]Sydney School of Public Health, The University of Sydney, Sydney, NSW, Australia, [2]Department of Nephrology and Dialysis, Medical University of Vienna, Vienna, Austria, [3]Department of Surgical Sciences, Sapienza University of Rome, Rome, Italy, [4]Department of Nephrology and Medical Intensive Care, Charité Universitätsmedizin Berlin, Berlin, Germany, [5]Department of Population Health Sciences, University of Bristol, Bristol, United Kingdom, [6]Department of Public Health and Primary Care, Academic Center for Nursing and Midwifery, KU Leuven, Leuven, Belgium, [7]Centre for Evidence in Transplantation, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, United Kingdom, [8]Department of Nephrology, Dialysis and Transplantation, Toulouse University Hospital, Toulouse, France, [9]Department of General, Transplant and Thoracic Surgery, Medical University of Innsbruck, Innsbruck, Austria, [10]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

Patient-reported outcomes (PROs) that assess individuals' perceptions of life participation, medication adherence, disease symptoms, and therapy side effects are extremely relevant in the context of kidney transplantation. All PROs are potentially suitable as primary or secondary endpoints in interventional trials that aim to improve outcomes for transplant recipients. Using PRO measures (PROMs) in clinical trials facilitates assessment of the patient's perspective of their health, but few measures have been developed and evaluated in kidney transplant recipients; robust methodologies, which use validated instruments and established frameworks for reporting, are essential. Establishing a core PROM for life participation in kidney transplant recipients is a critically important need, which is being developed and validated by the Standardized Outcomes in Nephrology (SONG)-Tx Initiative. Measures involving electronic medication packaging and smart technologies are gaining traction for monitoring adherence, and could provide more robust information than questionnaires, interviews, and scales. This article summarizes information on PROs and PROMs that was included in a Broad Scientific Advice request on clinical trial design and endpoints in kidney transplantation. This request was submitted to the European Medicines Agency (EMA) by the European Society for Organ Transplantation in 2016. Following modifications, the EMA provided its recommendations in late 2020.

**Keywords: patient-reported outcome measure (PROM), patient perspective, adherence, life participation, SONG-Tx, PROMIS®**

OPEN ACCESS

***Correspondence:***
*Maarten Naesens*
*maarten.naesens@kuleuven.be*

*[†]These authors have contributed equally to this work*

***Received:*** *21 October 2021*
***Accepted:*** *17 February 2022*
***Published:*** *20 May 2022*

***Citation:***
*Tong A, Oberbauer R, Bellini MI, Budde K, Caskey FJ, Dobbels F, Pengel L, Rostaing L, Schneeberger S and Naesens M (2022) Patient-Reported Outcomes as Endpoints in Clinical Trials of Kidney Transplantation Interventions. Transpl Int 35:10134. doi: 10.3389/ti.2022.10134*

## INTRODUCTION

The importance of the patient's perspective on their own health in the assessment of benefits and risks of therapeutic interventions is widely acknowledged (1). Such information could be relevant for drawing regulatory conclusions regarding treatment effects, benefit/risk balance assessments, or specific therapeutic claims (2). A patient-reported outcome (PRO) describes information assessed

and reported directly by the individual about how they feel or function in relation to their health or treatment, without interpretation or modification by anyone else, including clinicians and researchers (1, 3). Examples of PROs include health-related quality of life (HRQoL), physical function, ability to work, specific symptoms related to the disease or its treatment (e.g., pain, fatigue, side effects), and treatment adherence. A PRO measure (PROM) is a standardized quantitative assessment that captures the impact of disease and treatment as perceived by the individual.

In clinical research, PROs may be used as primary, co-primary, secondary, or exploratory endpoints (1, 4). However, evidence for the psychometric robustness of PROMs is an important consideration for the selection of PROs as endpoints in trials. The European Medicines Association (EMA) guideline on clinical investigation of medicinal products for the treatment of rheumatoid arthritis, for instance, recommends that several PROs are considered for secondary or supportive endpoints (5). However, for most disease areas, PROs are rarely incorporated in drug labeling claims. For example, of 60 PRO claims in orphan drug applications approved by the EMA between 2012 and 2016, only 12 (21.7%) of the products contained PROs in clinical study sections of the Summary of Product Characteristics (SmPC) (6). In 12 SmPCs, PROMs were based on symptoms; five also utilized patient functioning. HRQoL-related claims were included in eight approvals. A PRO was the primary endpoint in SmPCs in four (31%), a secondary endpoint in eight (62%), and a tertiary endpoint in one of the 13 approvals with a PRO claim. PROs that were primary endpoints assessed disease-specific symptoms exclusively (6).

Likewise, PROs are infrequently reported in kidney transplantation trials. Although regulatory agencies increasingly support the inclusion of PROs in clinical trials, few studies of medication regimens in kidney transplantation conform to these recommendations. One systematic review, for example, reported that only 2% of maintenance immunosuppression studies in kidney transplantation reported HRQoL outcomes (7). Another systematic review of 397 trials involving 63,514 adult kidney transplant recipients found substantial variability in PROs being assessed, as well as in PROMs used; the most frequent PROs were pain (40 trials, 15 measures), adherence (15 trials, eight measures), sleep (11 trials, four measures), and fatigue (11 trials, five measures) (8). Heterogeneity in choice of PROMs makes it difficult to compare intervention effects across trials. The PRO Rosetta Stone project developed and applied methods to link the patient-reported outcomes measurement information system (PROMIS) with other relevant measures, to provide equivalent scores for different scales that measure the same outcome (9). Also, there is limited evidence on the psychometric properties of PROMs used in kidney transplant recipients (10, 11).

This article provides an evidence-based and recipient-centered overview of PROs to be included as primary, secondary, or exploratory endpoints in clinical trials of kidney transplantation. Guidance on PRO measurement is also included, and the need for reliable measurement of medication adherence in randomized controlled trials (RCTs) is discussed.

## PROS TO BE INCLUDED IN RCTS INVOLVING KIDNEY TRANSPLANT RECIPIENTS

Two sections of the EMA's CHMP guideline on clinical investigation of immunosuppressants for solid organ transplantation (11) refer briefly to the incorporation of PROs in RCTs. Section 4.3.2 (definition of secondary endpoints) mentions HRQoL in the list of other frequently reported endpoints that can be included, yet does not consider HRQoL as a mandatory primary or secondary outcome of RCTs within transplant recipient populations. Section 4.4.3b (therapeutic studies; confirmatory trials) mentions adherence in the first aim of product development based on comparative trials, namely, "to substitute one or several therapeutic components of well-established immunosuppressive regimens to improve efficacy, safety or compliance" (12).

Selecting the right PRO involves identifying outcomes that are important to individuals, in addition to what might be relevant to the study hypothesis and intervention. Based on consensus from transplant recipients, caregivers, and healthcare professionals (HCPs), the following PROs could be considered to be incorporated in RCTs of kidney transplantation interventions as primary or secondary outcomes: life participation; medication adherence; symptoms and side effects.

### Life Participation

Through a consensus process involving over 1100 recipients, caregivers, and HCPs from 79 countries, the global Standardized Outcomes in Nephrology (SONG)-Tx initiative has established six core outcomes that should be reported in all kidney transplantation trials (**Figure 1**) (7, 8, 13). Alongside clinical outcomes relating to allograft loss, cardiovascular disease and mortality, cancer, and infection, life participation was the PRO of greatest importance to recipients, caregivers, and HCPs. Life participation describes "the ability to participate in activities that give patients a sense of fulfillment, enjoyment, control and hope in their lives" (14). Patients prefer not to specify life activities as these differ among individuals, so using a generic term enables life participation to be interpreted based on their own context (14).

### Medication Adherence

Non-adherence to prescribed medication is a global health concern. Adherence is divided into three quantifiable phases (15):

- Initiation (whether a patient takes the first prescribed dose)
- Implementation (the extent to which a patient's actual dosing corresponds to the prescribed regimen, from initiation until last dose taken)
- Discontinuation (when no more doses are taken, with persistence indicating length of time between initiation and last dose taken).

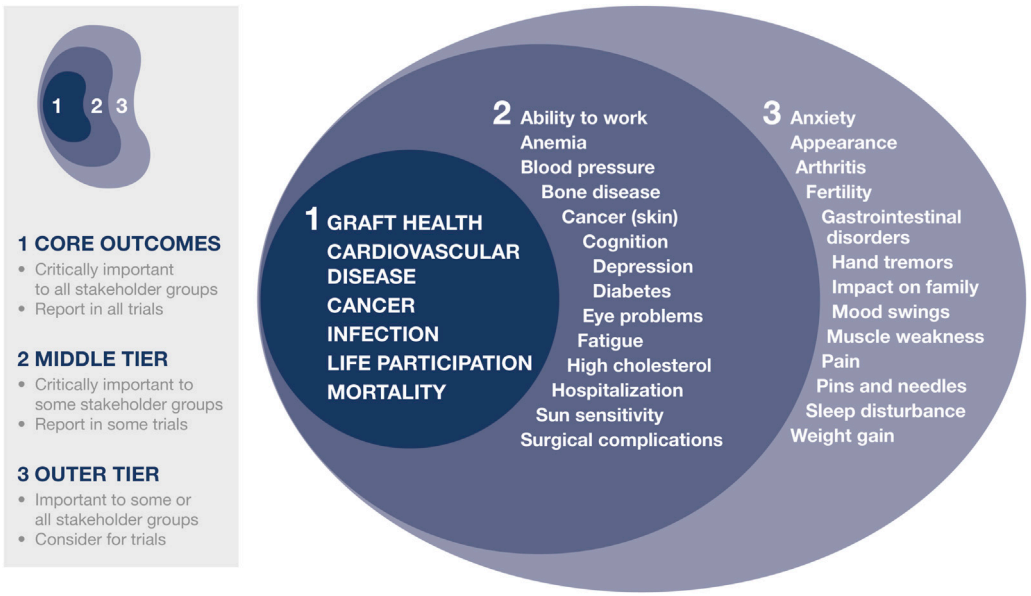


**FIGURE 1 |** SONG-Tx core outcomes. Reprinted from Kidney Int. Vol 94. Tong A, et al. (13). Implementing core outcomes in kidney disease. Report of the Standardized Outcomes in Nephrology, with permission from Elsevier.

## Concerns Associated With Non-adherence

Annually, across Europe, medication non-adherence contributes to ~200,000 avoidable deaths, and costs around €125 billion for excess hospitalizations, emergency care, and outpatient visits (16). Because of its impact on people's health, the EU advocates improving medication adherence as a key policy lever to minimize waste and optimize value derived from pharmaceutical expenditure (17).

Medication non-adherence is a major concern in solid organ transplantation (18). To maintain allograft function, recipients are prescribed complex regimens, typically including immunosuppressants and drugs to prevent or treat comorbidities. On average, following kidney transplantation, recipients take 22 pills daily (range, 8–47) at 3 months, and 23 pills daily (range, 9–57) at 12 months (19), with ~30% of the pill burden attributed to immunosuppressants (19, 20).

Compared with other solid organ transplant groups, kidney transplant recipients are the most vulnerable to non-adherence, with implementation problems occurring far more frequently than treatment discontinuation. Annually, over one-third of transplant recipients struggle to implement immuno-suppressive regimens correctly (21); deviations are commonly missed doses, incorrect dosing, or suboptimal timing of intake (22). Evidence consistently shows that poor implementation of an immunosuppressive regimen is an independent risk factor for rejection and allograft loss (23, 24). In addition, minor deviations from the regimen increase the risk of poor clinical outcome because of the narrow therapeutic window that exists for many immunosuppressant drugs (25).

The FDA supports the collection, analysis, and integration of patient perspective in the development of medical products and devices (26, 27). As part of their patient-focused drug development initiative, the US Food and Drug Administration (FDA) met with solid organ transplant recipients, caregivers, and advocates to elicit perceptions relating to recipients' well-being and treatment (2). Participants deemed medication adherence to be important, yet strict regimens posed challenges because of the frequency and high quantities of drugs, the need for clinic visits to monitor allograft function, the impact of therapy side effects, and difficulty remembering to take medications. Participants expressed a need for therapies that maintain long-term organ function, have fewer long-term comorbidities (such as cancer), have fewer side effects, and offer reduced frequency of administration compared with standard of care (28). Besides simplifying regimens and reducing symptom burden, patients wanted individualized treatment.

Another FDA-convened open public workshop on antibody-mediated rejection in kidney transplantation, which involved participants from academia and industry in addition to transplant recipients, also concluded that the prevalence of non-adherence is high and must be addressed, to improve transplant outcomes (29).

## Problems Associated With Assessing Adherence

When testing competing modes of drug treatment, it is essential to know the level of adherence to the regimen, including timely initiation, and punctual and sustained implementation, throughout the study.

Most deviations from a prescribed regimen can remain unnoticed yet jeopardize efficacy, safety, and selection of optimal dosing (15, 30–33). The gap between prescribed and actual drug-dosing history increases the risk of type II errors, as

the combined effects of variable underdosing and increased variance in response weaken statistical power for any demonstration of efficacy (34).

Non-adherence might also result in higher doses being prescribed, to achieve target trough levels, which could increase the risk of toxicity in adherent patients (35).

Regulatory agencies acknowledge the importance of assessing adherence. In its industry guidance on RCTs to support drug approval and biological products for human use, the FDA recommends identifying and selecting transplantation candidates who are likely to adhere to the regimen, and advocates quantification of adherence throughout a study (36). Regulation 536/2014 of the European Parliament and the Council on Clinical Trials on Medicinal Products for Human Use also stipulates that the initial application dossier should include "a description of procedures for monitoring subject compliance, if applicable" (37). In Europe, the EMA also published the ICH E9 (R1) addendum on estimands and sensitivity analysis in RCTs to the guideline on statistical principles for clinical trials, requesting researchers to consider adherence when quantifying treatment effects (38).

Unfortunately, despite regulatory guidance, adherence is rarely given prominence in RCTs. Suboptimal measures continue to be used, as regulatory agencies provide no or limited guidelines on how adherence should be assessed. Although SPIRIT guidelines describe strategies to be applied within RCTs (to improve adherence to intervention protocols, and procedures for monitoring adherence) (39), SPIRIT is vague on how adherence is best assessed, and only provides examples of suboptimal measures (e.g., tablet return). We advocate reliable, quantifiable methods for adherence measurement in kidney transplantation later in this article.

### Patient-Reported Symptoms and Side Effects

The SONG-Tx initiative (13, 40) and the FDA meeting on patient-focused drug development and adherence (28) revealed that transplant recipients are concerned about the number and burden of side effects associated with immunosuppression. These include the onset of serious comorbidities and debilitating symptoms, such as fatigue or pain (**Figure 1**).

In RCTs, side effects are typically assessed by adverse event checklists, completed by the treating physician. Although adverse event reporting is vitally important to monitor safety, empirical evidence indicated that adverse event checklists identified only 7% of symptoms experienced by patients (30). A systematic review of adverse event reporting in 233 trials of maintenance immunosuppression following kidney transplantation found inadequacies including selective reporting, poor definition and description of measurement, and lack of alignment with known and common side effects (41). Consequently, in transplantation studies, the true burden of immunosuppressive regimens remains underestimated, in terms of the number and severity of adverse events, and the overall distress associated with treatment-related symptoms. Individuals may find it difficult to determine whether their symptoms relate to medications or their health condition (28), but irrespective of underlying causes, side effects and symptoms are important determinants of HRQoL, and might trigger non-adherence (42). PROMs can support patients in expressing how they feel and function so that, in turn, clinicians can aim to better manage patients' symptoms (and how these impact on life), to improve patient-centered care. Therefore, we recommend that patient-reported symptoms and side effects represents meaningful primary or secondary endpoints for RCTs in kidney transplantation.

## SELECTION OF APPROPRIATE PROMS

Frameworks are available to guide the selection of PROMs for use in RCTs (3, 43–45). The rationale for selecting PROMs for a RCT should consider the prevalence and nature of the condition, characteristics that are relevant or unique to the patient population, patient perspectives and priorities, and outcomes that might be expected to change in response to the intervention (3, 45, 46).

PROMs can be classified into one of three categories (44). Firstly, there are generic health status measures, which assess a range of constructs [usually a combination of impairment, disability, and HRQoL (46)]. These can apply across different conditions or populations, are useful for broad comparisons of the relative impact of interventions between diseases, and can be compared with population normative data. Such measures include the 36-item Short Form Health Survey (SF-36), the World Health Organization Quality of Life Scale (WHO-QOL), and PROMIS®-29 (47). PROMIS-29 (and PROMIS-57) profile instruments that include the ability to participate in social roles and activities scale, and both have been validated in kidney transplant recipients (48, 49). Secondly, condition- or symptom-specific measures assess PROs within either a condition or disease, or across certain symptoms. Examples include the Kidney Disease Quality of Life instrument (KDQoL) (50), Kidney Transplant Questionnaire (51), Modified Transplant Symptom Occurrence and Symptom Distress Scale (52), End-stage Renal Disease Symptom Checklist—Transplantation Module (ESRD-SCLTM) (53), and Gastrointestinal Symptom Rating Scale (GSRS) (54). Finally, preference-based (or utility) measures assess a value (i.e., from <0 [worse than being dead] to 1 [full health]), assigned to the health state described by the patient. Values are assessed using direct methods (such as time trade-off or standard gamble), or multi-attribute utility instruments (55). A utility value allows comparison of HRQoL across conditions and between populations. In economic evaluations, these measures can be used to calculate quality-adjusted life-years (QALYs) to provide cost-effectiveness findings. Examples of preference-based measures include the EQ-5D, Health Utilities Index, and time trade-off calculations. In economic evaluations, such measures can be used to calculate QALYs and in doing so provide cost-effectiveness findings. Examples of preference-based measures are EQ-5D, HUI, and time trade-off. Data from the KDQoL/SF-36 and PROMIS profile measures can also be used in economic evaluations.

The UK Health Technology Assessment Programme recommends eight criteria for PROM selection: appropri-ateness, reliability (internal consistency, reproducibility), validity (criterion and predictive validity, face and content validity, construct validity), responsiveness, precision, interpretability, acceptability, and feasibility (43). COMET guidelines can be used to develop a core outcome set (COS), defined as a minimum set of outcomes that should be reported in all studies within a specific condition or population (56); in addition, the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) initiative provides specific recommendations for selecting the most appropriate measures of COS (57) (https://www.cosmin.nl/). Core PROMs have been identified in kidney disease (e.g., fatigue in people undergoing hemodialysis) (58) and for other health conditions. The COMET and OMERACT initiatives recommend that a core outcome set includes a PRO (59) and there are frameworks for selecting core outcome measures for PROs (56, 59). Of note, the FDA has released guidance for core PROs in clinical trials in oncology (60). Below, we suggest measures to assess core outcomes for RCTs in kidney transplantation.

## LIFE PARTICIPATION AS A CORE OUTCOME MEASURE

Following a consensus workshop on establishing a core outcome measure for life participation, kidney transplant recipients, caregivers, and HCPs recommended that such a measure needs to achieve several milestones. Firstly, it should capture recipients' goals to fulfill their roles and re-establish a normal lifestyle post-transplantation. It should also include the diverse activities of "life" as defined by recipients, capture life changes caused by treatment complications and side effects, and be validated and feasible to implement (14).

A systematic review of 230 trials and observational studies found that 29 measures have been used to assess life participation in kidney transplant recipients (61). The most frequently used were the SF-36, KDQoL, and EQ-5D, which capture aspects of life participation in one attribute, although few instruments specifically measured aspects of life participation. Validation data were available for only six measures, and no validation data were available for the subscale capturing life participation. Also, none of the instruments adequately addressed recipients' perspectives and experiences of life participation (14, 61). Establishing a core PROM for life participation in kidney transplantation populations is therefore needed, to ensure consistent reporting of this critically important outcome.

The SONG-Tx initiative suggested that PROMIS SF v2.0, *Ability to participate in social roles and activities*, was the best available measure to capture transplant recipients' perspectives, priorities, and experiences regarding life participation (62). PROMIS items are available in ~30 languages and have been rigorously validated (63); exploratory factor analysis, confirmatory factor analysis, item response theory modeling, and evaluation of differential item functioning were also used to test items (64). Cross-sectional evidence supports the validity of PROMIS items, and the reliability and precision of generic symptoms and functional reports: findings for PROMIS are comparable with other well-validated and widely accepted measures (65, 66).

Evidence also supports the psychometric robustness of PROMIS SF v2.0: sufficient unidimensionality, local dependence, monotonicity, graded response model item fit, and differential item functioning for age, sex, education, region, ethnicity, and language were demonstrated in a Dutch population (N = 1002) (67). Reliability, and content and construct validity, have been shown for PROMIS instruments (including ability to participate in social roles) in people with rheumatoid arthritis and cancer, and in clinical care settings (68).

The PROMIS measure was adapted by SONG-Tx following cognitive interviews with kidney transplant recipients [N = 20]. These were conducted using a pre-testing framework based on cognitive and social psychology, which assessed aspects of respondents' comprehension, retrieval, response, and judgment (69). Initial findings indicated that kidney transplant recipients preferred positive wording compared with the focus on "trouble" used in the original PROMIS measure. In addition, if life participation is a primary outcome, use of a long measure is recommended, to facilitate comprehensive assessment (62). Following this preliminary work, items were adapted based on extensive input from kidney transplant recipients, with modifications being reviewed before generating the final Song-Tx Life Participation measure. A validation study in kidney transplantation is in progress and will be completed before recommending its use as a core outcome measure. More information on the Song-Tx Life Participation measure is available from the authors on request.

Other measures of health status or instruments assessing variables that might influence life participation (e.g., depression, fatigue) may also be required to address the specific aims of a study and/or intervention. For example, RCTs commonly include an economic evaluation to demonstrate cost-effectiveness, usually based on the benefits of the intervention, measured as QALYs. In this case, a utility-based instrument such as EQ-5D would be required, but this should be in addition to—not in place of—measuring life participation. We recommend that selecting additional PROMs should be undertaken in accordance with COSMIN guidelines or its equivalent.

## MEASURES OF MEDICATION ADHERENCE

The COMMIT clinician group recommends measurement of medication adherence as the "fifth vital sign" in transplantation studies (18). Choice of method depends on phase of adherence under investigation (initiation, implementation, discontinuation), context of use (RCT, routine care, or registry), study purpose (observational or interventional), reliability and richness of data sought, participants' preferences, and usability of the measures (70).

A systematic review of studies involving various chronically ill patient populations identified 20 different

self-report measures for capturing non-adherence (71). The Basel Assessment of Adherence to Immunosuppressive Medication Se (BAASIS®; http://baasis.nursing.unibas.ch/) (72) was recommended for assessing adherence to immunosuppressive drug regimens because it is short to perform and easy to score, focuses on the implementation phase, considers both taking and timing of intake, and has established reliability and validity (71, 72). COMMIT also recommends using the BAASIS®, alongside the Insulin Treatment Appraisal Scale and Simplified Medication Adherence Questionnaire (18). However, more research on the prognostic value of self-reporting is needed, as PROMs might underestimate medication non-adherence.

Unlike self-report measures, smart technology allows for continuous measurement of adherence behaviors, providing objective data on day-to-day variability and timing of taking medication. Identifying the timing of gaps in drug exposure is needed when aiming to develop a reliable efficacy and safety profile. Smart technology is also recommended by the FDA to improve adherence measurement in RCTs (36). There are three broad categories of methods to measure adherence by means of smart technology (35). Firstly, video-assisted or photographic documentation of drug intake (e.g., using a mobile phone app with face-recognition technology, capturing the patient taking medication). Secondly, electronic detection of package entry (by incorporating a microchip in the container or registering and time-stamping removal of a single dose: the latest systems allow for real-time transfer of information on dosing and timing to the researcher). Finally, ingestible smart sensors, embedded in pharmaceuticals, which send a time-stamped signal (activated by gastric acid) to a patch worn on the patient's skin (known as "raisin technology"). Additional research is required, to determine the accuracy, usability, and acceptability of smart technology before their use in drug trials can be recommended.

The use of electronic monitoring devices could be considered, following positive experiences in solid organ transplantation: a multicenter RCT employed such technology successfully in 219 participants to compare medication adherence (primary endpoint) between modified-release tacrolimus once-daily and twice-daily regimens (22). The feasibility of such monitoring in kidney transplant populations could be justified, relative to the drug-development and overall costs of RCTs; return on investment could be substantial, given that it is the only method available to visualize daily drug-intake patterns. However, self-reporting of medication adherence should be embedded in all clinical trials, irrespective whether such PROMs can be combined with smart technology measures.

## PROMS FOR SYMPTOMS AND SIDE EFFECTS

Instruments to measure the impact of symptoms and side effects should be selected based on similar criteria relating to reliability, validity, and responsiveness to change as those proposed by COSMIN (57). The instrument chosen should be determined by the RCT aims and the type of intervention. For example, an intervention that aims to alleviate gastrointestinal side effects—either through additional medication or substitution of immunosuppressants or dose adjustments—might select the GSRS and the Gastrointestinal Quality of Life Index as being relevant validated instruments (54). Similarly, instruments assessing anxiety, depression, or mood swings may be required to address the intervention aim.

Instruments can also report frequency and severity of side-effect profiles associated with immunosuppression. A systematic review applying the COSMIN checklist to appraise the psychometric quality of PROMs used in patients with kidney disease deemed the ESRD-SCLTM to be the most suitable measure for use in research and clinical practice, as it had strong evidence for internal consistency, and moderate evidence for test/retest reliability and structural and construct validity (10). However, the authors of this review also noted that no instrument had evidence supporting all measurement properties.

## PRO REPORTING

We recommend that PRO reporting in study protocols should follow the international, consensus-based SPIRIT-PRO extension guidelines, with CONSORT-PRO used for reporting RCT results (4).

## CONCLUSIONS

- The use of PROMs in RCTs enables assessment of the patient's perspective of their own health:
  - PROM selection requires consideration of the appropriateness, reliability, validity, acceptability, and feasibility of use.
  - Transparent reporting on the use and results of PROMs, using established frameworks, is required.
- The PROs life participation, medication adherence, and symptoms and side effects are suitable secondary endpoints in interventional studies.
  - These PROs are relevant and important in the context of kidney transplantation.
- Electronic monitoring to document adherence in RCTs is advised.
  - If this is not feasible, self-report measures such as the BAASIS might be considered, bearing in mind that self-reporting data has limited reliability and does not capture day-to-day patterns of medication intake or regularity of intake.
- SPIRIT-PRO should be used for reporting study protocols, and CONSORT-PRO for reporting RCT results.
- Physical, emotional, and cognitive functioning, mental health, and health-related quality of life are relevant PRO domains to be considered in trials in kidney transplantation.

## Scientific Advice from the Committee for Medicinal Products for Human Use of the European Medicines Agency Regarding These Conclusions

- CHMP acknowledged that PRO measures are important to capture the patient's perception; nevertheless, assessment of PRO data is difficult due to the nature of such data.
  - The benefit/risk assessment of clinical trials addresses many of the issues participants express as being important, which include achieving long-term organ function with fewer comorbidities or adverse events.
- Among the most frequently cited instruments to address generic health status are the SF-36, the Sickness Impact Profile, and the WHO-QOL.
  - ESOT suggests how to establish a "core outcome set" in clinical trials of transplantation that do not incorporate instruments frequently used to address generic health status. These core outcomes should:
    - Capture recipients' goals to fulfill their roles and re-establish a normal lifestyle
    - Include the diverse activities of "life" as defined by recipients
    - Capture life changes caused by complications and side effects of treatment
    - Be validated and feasible to implement (14).
- CHMP supported the inclusion of HRQoL measures and agreed that a validated PRO tool could be important.
  - PROs are often included as secondary measures of efficacy in clinical trials.
  - Use of a PRO as primary endpoint would require predefining a clinically meaningful improvement as measured by the PRO and powering of the study to this difference; the CHMP was not aware of a consensus defining such difference.
  - There is a burden in participating in clinical trial for each patient, generally higher than in normal clinical practice.
- The CHMP agreed that the selection of PROs requires consideration of the appropriateness, reliability, validity, acceptability, and feasibility of use, without causing excess burden to the study participant.
  - The proposed SPIRIT-PRO extension guideline and use of the CONSORT-PRO for reporting the results of randomized trials (4) are acceptable; guidelines on reporting files from clinical trials are provided in GCP guideline EMA/INS/GCP/856758/2018.
- The CHMP agreed that medication adherence should be measured by reliable methods in clinical trials, considering ICH E9 (R1) addendum on estimands and sensitivity analyses should be performed.
  - The count of returned tablets is not deemed a fully reliable measure of adherence, but it is an important tool used in clinical trials.
  - Medication adherence is known to be better during clinical trials and to decrease over the course of regular treatment, especially if the treatment is lifelong; therefore, medication adherence is an important PRO to evaluate.
    - However, the assessment of medication adherence in a clinical trial could prove to be difficult to extrapolate to real life.
  - Measures involving electronic medication packaging and different smart technologies are gaining traction for measuring adherence and could provide more robust information than questionnaires, interviews, and scales.

## AUTHOR CONTRIBUTIONS

This article is one of many developed from the Broad Scientific Advice request, submitted to the EMA/CHMP by ESOT in 2020: interactions between the EMA and ESOT regarding this request began in 2016. For the present article, the working group on patient-reported outcomes developed the ESOT position on the core question "Does CHMP agree with the proposed patient-reported outcomes as (primary/secondary) endpoints for use in clinical trials of kidney transplantation interventions?" The Centre for Evidence in Transplantation provided support with specific data extraction requests: these literature searches formed the basis of evidence used in the advice request and the present article. Input into the working group's output was provided from all ESOT members involved in the advice request process. The present article was adapted by MN from the final Broad Scientific Advice request submission (June 2020), presentation documents and minutes of the meeting between ESOT and the CHMP Scientific Advice Working Party (SAWP) (September 2020), and the final response from the SAWP (December 2020). The first draft of the article was reviewed by AT and RO; the revised draft was reviewed, finalized, and approved by all co-authors before submission.

## FUNDING

This initiative was supported by the European Society for Organ Transplantation.

## CONFLICT OF INTEREST

RO has received grants/research support from Amgen, Astellas, and Chiesi; and speakers' bureaux/honoraria from Amgen, Astellas, Chiesi, Hansa, Neovii, Novartis, and Teva. KB has received honoraria and/or research funding from Alexion, Astellas, Bristol Myers Squibb, Chiesi, Fresenius, Hansa, Hexal, Merck, Novartis, Otsuka, Pfizer, Roche, Sandoz, Siemens, and Veloxis. SS has received grants/research support from Bridge to Life, Chiesi, Neovii, Novartis, Organ Recovery Systems, and Sandoz; speakers' bureaux/honoraria from Astellas, BMS, Chiesi, Novartis, OrganOx, and Sanofi; and consulting fees from Astellas, Atara, Merck, NefroHealth, Novartis, Sandoz, and Teva.

The remaining authors declare that the work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## ACKNOWLEDGMENTS

The authors thank the experts involved with the European Medicines Agency (EMA) Committee for Medicinal Products for Human Use (CHMP) who participated in the Broad Scientific Advice request. The published information is based on EMA feedback received during the Broad Scientific Advice request. EMA/CHMP have not been involved in the drafting or review of the manuscript to be published. This publication does not constitute a formal EMA/CHMP endorsement of the manuscript. Medical writing support was provided by Linda Edmondson, independent medical writer, funded by ESOT.

## REFERENCES

1. Bansal D, Bhagat A, Schifano F, Gudala K. Role of Patient-Reported Outcomes and Other Efficacy Endpoints in the Drug Approval Process in Europe (2008-2012). *J. Epidemiol Global Health* (2015). 5:385–95. doi:10.1016/j.jegh.2015.04.006
2. Committee for Medicinal Products for Human Use (CHMP). EMEA/CHMP/EWP/139391/2004. Reflection Paper on the Regulatory Guidance for the Use of Health-Related Quality of Life (HRQL) Measures in the Evaluation of Medicinal Products (2022). Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/reflection-paper-regulatory-guidance-use-healthrelated-quality-life-hrql-measures-evaluation_en.pdf (Accessed January 31, 2022).
3. U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims: Draft Guidance. *Health Qual Life Outcomes* (2006). 4:79. doi:10.1186/1477-7525-4-79
4. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan A-W, King MT. The SPIRIT-PRO Group. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. *JAMA* (2018). 319:483–94. doi:10.1001/jama.2017.21903
5. Committee for Medicinal Products for Human Use. Guideline on Clinical Investigation of Medicinal Products for the Treatment of Rheumatoid Arthritis (CPMP/EWP/556/95) (2022). Available at: https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-clinical-investigation-medicinal-products-treatment-rheumatoid-arthritis_en.pdf (Accessed January 31, 2022).
6. Jaroslawski S, Ququier P, Borissov B, Dussart C, Toumi M. Patient-reported Outcome Claims in European and United States Orphan Drug Approvals. *J Mark Access Health Pol* (2018). 6:1542920.
7. Howell M, Wong G, Turner RM, Tan HT, Tong A, Craig JC, et al. The Consistency and Reporting of Quality of Life Outcomes in Trials of Immunosuppressive Agents in Kidney Transplantation: A Systematic Review and Meta-Analysis. *Am J Kidney Dis* (2016). 67:762–74. doi:10.1053/j.ajkd.2015.11.006
8. Sautenet B, Chapman TAJR, Warrens AN, Rosenbloom D, Wong G, et al. Range and Consistency of Outcomes Reported in Randomized Trials Conducted in Kidney Transplant Recipients: A Systematic Review. *Transplantation* (2018). 102:2065–71. doi:10.1097/tp.0000000000002278
9. PROsetta Stone (2022). Available at: https://www.prosettastone.org (Accessed January 31, 2022).
10. Aiyegbusi O, Kyte D, Cockwell P, Marshall T, Gheorghe A, Keeley T, et al. Measurement Properties of Patient-Reported Outcome Measures (PROMs) Used in Adult Patients with Chronic Kidney Disease: A Systematic Review. *PLoS One* (2017). 21:e0179733. doi:10.1371/journal.pone.0179733
11. Cleemput I, Kesteloot K, Moons P, Vanrenterghem Y, Van Hooff JP, Squifflet JP, et al. The Construct and Concurrent Validity of the EQ-5D in a Renal Transplant Population. *Value Health* (2004). 7:499–509. doi:10.1111/j.1524-4733.2004.74013.x
12. Committee for Medicinal Products for Human Use. Guideline on Clinical Investigation of Immunosuppressants for Solid Organ Transplantation (CHMP/EWP/263148/06) (2008). Available at: https://www.ema.europa.eu/en/clinical-investigation-immunosuppressants-solid-organ-transplantation (Accessed January 31, 2022).
13. Tong A, Manns B, Wang AYM, Hemmelgarn B, Wheeler DC, Gill J, et al. Implementing Core Outcomes in Kidney Disease: Report of the Standardized Outcomes in Nephrology (SONG) Implementation Workshop. *Kidney Int* (2018). 94:1053–68. doi:10.1016/j.kint.2018.08.018
14. Ju A, Josephson MA, Butt Z, Jowsey-Gregoire S, Tan J, Taylor Q, et al. Establishing a Core Outcome Measure for Life Participation: A Standardized Outcomes in Nephrology–Kidney Transplantation Consensus Workshop Report. *Transplantation* (2018). 103:1199–205. doi:10.1097/TP.0000000000002476
15. Vrijens B, De Geest S, Hughes DA, Przemyslaw K, Demonceau J, Ruppar T, et al. A New Taxonomy for Describing and Defining Adherence to Medications. *Br J Clin Pharmacol* (2012). 73:691–705. doi:10.1111/j.1365-2125.2012.04167.x
16. Khan R, Socha-Dietrich K. Investing in Medication Adherence Improves Health Outcomes and Health System Efficiency: Adherence to Medicines for Diabetes, Hypertension, and hyperlipidaemia. OECD Health Working Papers (2018). Available at: https://www.oecd-ilibrary.org/social-issues-migration-health/investing-in-medication-adherence-improves-health-outcomes-and-health-system-efficiency_8178962c-en (Accessed January 31, 2022).
17. OECD/European Union. *Health at a Glance: Europe 2018: State of Health in the EU Cycle*. Paris: OECD Publishing (2018) (Accessed January 31, 2022).
18. Neuberger JM, Bechstein WO, Kuypers DR, Burra P, Citterio F, De Geest S, et al. Practical Recommendations for Long-Term Management of Modifiable Risks in Kidney and Liver Transplant Recipients: A Guidance Report and Clinical Checklist by the Consensus on Managing Modifiable Risk in Transplantation (COMMIT) Group. *Transplantation* (2017). 101(4S Suppl. 2):S1. doi:10.1097/TP.0000000000001651
19. Low JK, Crawford K, Manias E, Williams A. Quantifying the Medication burden of Kidney Transplant Recipients in the First Year Post-transplantation. *Int J Clin Pharm* (2018). 40:1242–9. doi:10.1007/s11096-018-0678-9
20. Hardinger KL, Hutcherson T, Preston D, Murillo D. Influence of Pill Burden and Drug Cost on Renal Function after Transplantation. *Pharmacotherapy* (2012). 32:427–32. doi:10.1002/j.1875-9114.2012.01032.x
21. Dew MA, Dimartini AF, De Vito Dabbs A, Myaskovsky L, et al. Rates and Risk Factors for Nonadherence to the Medical Regimen after Adult Solid Organ Transplantation. *Transplantation* (2007). 83:858–73. doi:10.1097/01.tp.0000258599.65257.a6
22. Kuypers DR, Peeters PC, Sennesael JJ, Kianda MN, Vrijens B, Kristanto P, et al. Improved Adherence to Tacrolimus Once-Daily Formulation in Renal Recipients: A Randomized Controlled Trial Using Electronic Monitoring. *Transplantation* (2013). 95:333–40. doi:10.1097/tp.0b013e3182725532
23. Davis S, Cooper JE. Acute Antibody-Mediated Rejection in Kidney Transplant Recipients. *Transpl Rev* (2017). 31:47–54. doi:10.1016/j.trre.2016.10.004
24. Al-Sheyyab A, Binari L, Shwetar M, Ramos E, Kapp ME, Bala S, et al. Association of Medication Non-adherence with Short-Term Allograft Loss after the Treatment of Severe Acute Kidney Transplant Rejection. *BMC Nephrol* (2019). 20:373. doi:10.1186/s12882-019-1563-z
25. Fine RN, Becker Y, De Geest S, Eisen H, Ettenger R, Evans R, et al. Nonadherence Consensus Conference Summary Report. *Am J Transpl* (2009). 9:35–41. doi:10.1111/j.1600-6143.2008.02495.x
26. US Food & Drug Administration. Principles for Selecting, Developing, Modifying, and Adapting Patient Reported Outcome Instruments for Use in Medical Device Evaluation. Guidance for Industry and Food and Drug

Administration Staff, and Other Stakeholders (2022). Available at: https://www.fda.gov/media/141565/download (Accessed January 31, 2022).
27. US Food & Drug Administration. Guidance for Industry. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims (2022). https://www.fda.gov/media/77832/download (Accessed January 31, 2022).
28. Ettenger R, Albrecht R, Alloway R, Belen O, Cavaillé-Coll MW, Chisholm-Burns MA, et al. Meeting Report: FDA Public Meeting on Patient-Focused Drug Development and Medication Adherence in Solid Organ Transplant Patients. *Am J Transpl* (2018). 18:564–73. doi:10.1111/ajt.14635
29. Velidedeoglu E, Cavaillé-Coll MW, Bala S, Belen OA, Wang Y, Albrecht R. Summary of 2017 FDA Public Workshop: Antibody-Mediated Rejection in Kidney Transplantation. *Transplantation* (2018). 102:e257–64. doi:10.1097/tp.0000000000002141
30. Anderson RB, Testa M. Symptom Distress Checklists as a Component of Quality of Life Measurement: Comparing Prompted Reports by Patient and Physician with Concurrent Adverse Event Reports via the Physician. *TIRS* (1994). 28:114. doi:10.1177/009286159402800112
31. Blaschke TF, Osterberg L, Vrijens B, Urquhart J. Adherence to Medications: Insights Arising from Studies on the Unreliable Link between Prescribed and Actual Drug Dosing History. *Ann Rev Pharmacol Toxicol* (2012). 52:275–301. doi:10.1146/annurev-pharmtox-011711-113247
32. Breckenridge A, Aronson JK, Blaschke TF, Hartman D, Peck CC, Vrijens B. Poor Medication Adherence in Clinical Trials: Consequences and Solutions. *Nat Rev Drug Discov* (2017). 16:149–50. doi:10.1038/nrd.2017.1
33. Shiovitz TM, Bain EE, McCann DJ, Skolnick P, Laughren T, Hanina A, et al. Mitigating the Effects of Nonadherence in Clinical Trials. *J Clin Pharmacol* (2016). 56:1151–64. doi:10.1002/jcph.689
34. Mallayasamy S, Chaturvedula A, Blaschke T, Fossler MJ. A Systematic Evaluation of Effect of Adherence Patterns on the Sample Size and Power of a Clinical Study. *CPT Pharmacometrics Syst Pharmacol* (2018). 7:818–28. doi:10.1002/psp4.12361
35. Vrijens B, Urquhart J. Methods for Measuring, Enhancing, and Accounting for Medication Adherence in Clinical Trials. *Clin Pharmacol Ther* (2014). 95:617–26. doi:10.1038/clpt.2014.59
36. US Food and Drug Administration. Guidance for Industry: Enrichment Strategies for Clinical Trials to Support Approval of Human Drugs and Biological Products (2019). Available at: https://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm332181.pdf (Accessed January 31, 2022).
37. European Union (EU). Official Journal of the European Union. Regulation (EU) No 536/2014 of the European Parliament and of the Council of 16 April 2014 on Clinical Trials on Medicinal Products for Human Use, and Repealing Directive 2001/20/EC (2014). Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014R0536&rid=1 (Accessed January 31, 2022).
38. European Medicines Agency. Conditional Marketing Authorization: Report on Ten Years of Experience at the EMA (2017). Available at: https://www.ema.europa.eu/en/documents/report/conditional-marketing-authorisation-report-ten-years-experience-european-medicines-agency_en.pdf (Accessed January 31, 2022).
39. Chan AW, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 Statement: Defining Standard Protocol Items for Clinical Trials. *Ann Intern Med* (2013). 158:200–7. doi:10.7326/0003-4819-158-3-201302050-00583
40. Tong A, Gill J, Budde K, Marson L, Reese PP, Rosenbloom D, et al. Toward Establishing Core Outcome Domains for Trials in Kidney Transplantation: Report of the Standardized Outcomes in Nephrology-Kidney Transplantation Consensus Workshops. *Transplantation* (2017). 101:1887–96. doi:10.1097/tp.0000000000001774
41. Howell M, Yeo R, Tong A, Craig JC, Howard K, Wong G. Completeness of Reporting of Adverse Events in Trials of Maintenance Immunosuppression in Kidney Transplantation: A Systematic Review. *Nephrol Dial Transpl* (2017). 32:1586–94. doi:10.1093/ndt/gfx216
42. Kugler C, Geyer S, Gottlieb J, Simon A, Haverich A, Dracup K. Symptom Experience after Solid Organ Transplantation. *J Psychosom Res* (2009). 66:101–10. doi:10.1016/j.jpsychores.2008.07.017
43. Fitzpatrick R, Davey C, Buxton M, Jones DR. Evaluating Patient-Based Outcome Measures for Use in Clinical Trials: A Review. *Health Technol Assess* (1998). 2:1–74. doi:10.3310/hta2140
44. Ju A, Tong A. Considerations and Challenges in Selecting Patient-Reported Outcome Measures for Clinical Trials in Nephrology. *Clin J Am Soc Nephrol* (2017). 12:1882–4. doi:10.2215/cjn.06300617
45. Snyder CF, Watson ME, Jackson JD, Cella D, Halyard MY. Myao/FDA Patient-Reported Outcomes Consensus Meeting Group. Patient-Reported Outcome Instrument Selection: Designing a Measurement Strategy. *Value Health* (2007). 10:S76–85. doi:10.1111/j.1524-4733.2007.00270.x
46. McKenna SP. Measuring Patient-Reported Outcomes: Moving Beyond Misplaced Common Sense to Hard Science. *BMC Med* (2011). 9:86. doi:10.1186/1741-7015-9-86
47. HealthMeasures. PROMIS®-29 (2021). Available at: https://www.healthmeasures.net/ (Accessed Octobeer 18, 2021).
48. Zhang J, Dewitt B, Tang E, Breitner D, Saqib M, Li D, et al. Evaluation of PROMIS Preference Scoring System (PROPr) in Patients Undergoing Hemodialysis or Kidney Transplant. *Clin J Am Soc Nephrol* (2021). 16:1328–36. doi:10.2215/cjn.01880221
49. Tang E, Ekundayo O, Peipert JD, Edwards N, Bansal A, Richardson C, et al. Validation of the Patient-Reported Outcomes Measurement Information System (PROMIS)-57 and -29 Item Short Forms Among Kidney Transplant Recipients. *Qual Life Res* (2019). 28:815–27. doi:10.1007/s11136-018-2058-2
50. Rand Healthcare. Kidney Disease Quality of Life Instrument (2021). Available at: https://www.rand.org/health-care/surveys_tools/kdqol.html (Accessed January 31, 2022).
51. Kidney Transplant Questionnaire (1993). Available at: https://eprovide.mapi-trust.org/instruments/kidney-transplant-questionnaire-25-items (Accessed January 31, 2022).
52. Dobbels F, Moons P, Abraham I, Larsen CP, Dupont L, De Geest S, et al. Measuring Symptom Experience of Side-Effects of Immunosuppressive Drugs: the Modified Transplant Symptom Occurrence and Distress Scale. *Transpl Int* (2008). 21:764–73. doi:10.1111/j.1432-2277.2008.00674.x
53. Franke GH, Reimer J, Kohnle M, Luetkes P, Maehner N, Heemann U, et al. Quality of Life in End-Stage Renal Disease Patients after Successful Kidney Transplantation: Development of the ESRD Symptom Checklist – Transplantation Module. *Nephron* (1999). 83:31–9. doi:10.1159/000045470
54. Chan S, Cao C, Pascoe EM, Johnson DW, Shah A, Holtmann GA, et al. Patient-reported Gastrointestinal Symptoms and the Association with Quality of Life Following Kidney Transplantation. *Kidney Int Rep* (2020). 6:138–45. doi:10.1016/j.ekir.2020.10.013
55. Wyld M, Morton RL, Hayen A, Howard K, Webster AC. A Systematic Review and Meta-Analysis of Utility-Based Quality of Life in Chronic Kidney Disease Treatments. *Plos Med* (2012). 9:e1001307. doi:10.1371/journal.pmed.1001307
56. Prinsen CA, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to Select Outcome Measurement Instruments for Outcomes Included in a 'Core Outcome Set' – a Practical Guideline. *Trials* (2016). 17:449. doi:10.1186/s13063-016-1555-2
57. Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) Initiative (2022). Available at: https://www.cosmin.nl/ (Accessed January 31, 2022).
58. Ju A, Teixeira-Pinto A, Tong A, Smith AC, Unruh M, Davison SN, et al. Validation of a Core Patient-Reported Outcome Measure for Fatigue in Patients Receiving Hemodialysis: The SONG-HD Fatigue Instrument. *Clin J Am Soc Nephrol* (2020). 15:1614–21. doi:10.2215/cjn.05880420
59. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: Version 1.0. *Trials* (2017). 18(Suppl. 3):280. doi:10.1186/s13063-017-1978-4
60. US Food & Drug Administration. Core Patient-Reported Outcomes in Cancer Clinical Trials Draft Guidance for Industry (2022). Available at: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/core-patient-reported-outcomes-cancer-clinical-trials (Accessed January 31, 2022).
61. Ju A, Chow J, Ralph AF, Howell M, Josephson MA, Ahn C, et al. Patient-reported Outcome Measures for Life Participation in Kidney Transplantation. *Am J Transpl* (2019). 19:2306–17. doi:10.1111/ajt.15267
62. PROMIS. PROMIS Short Form v2.0 – Ability to Participate in Social Roles and Activities 4a (2018). Available at: http://www.healthmeasures.

net/images/PROMIS/manuals/PROMIS_Ability_to_Participate_in_Social_Roles_and_Activities_Scoring_Manual.pdf (Accessed January 31, 2022).

63. Dewalt DA, Rothrock N, Yount S, Stone AA. PROMIS Cooperative Group. Evaluation of Item Candidates: The PROMIS Qualitative Item Review. *Med Care* (2007). 45:S12–21. doi:10.1097/01.mlr.0000254567.79743.e2
64. Hahn EA, Devellis RF, Bode RK, Garcia SF, Castel LD, Eisen SV, et al. Measuring Social Health in the Patient-Reported Outcomes Measurement Information System (PROMIS): Item Bank Development and Testing. *Qual Life Res* (2010). 19:1035–44. doi:10.1007/s11136-010-9654-0
65. Cella D, Riley W, Stone A, Rothrock N, Reeve B, Yount S, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) Developed and Tested its First Wave of Adult Self-Reported Health Outcome Item Banks: 2005–2008. *J Clin Epidemiol* (2010). 63:1179–94. doi:10.1016/j.jclinepi.2010.04.011
66. Cook KF, Jensen SE, Schalet BD, Beaumont JL, Amtmann D, Czajkowski S, et al. PROMIS Measures of Pain, Fatigue, Negative Affect, Physical Function, and Social Function Demonstrated Clinical Validity across a Range of Chronic Conditions. *J Clin Epidemiol* (2016). 73:89–102. doi:10.1016/j.jclinepi.2015.08.038
67. Terwee CB, Crins MHP, Boers M, de Vet HCW, Roorda LD. Validation of Two PROMIS Item Banks for Measuring Social Participation in the Dutch General Population. *Qual Life Res* (2019). 28:211–20. doi:10.1007/s11136-018-1995-0
68. Hahn EA, Kallen MA, Jensen RE, Potosky AL, Moinpour CM, Ramirez M, et al. Measuring Social Function in Diverse Cancer Populations: Evaluation of Measurement Equivalence of the Patient Reported Outcomes Measurement Information System® (PROMIS®) Ability to Participate in Social Roles and Activities Short Form. *Psychol Test Assess Model* (2016). 58:403–21.
69. Collins D. Pretesting Survey Instruments: An Overview of Cognitive Methods. *Qual Life Res* (2003). 12:229–38. doi:10.1023/a:1023254226592
70. Whalley Buono E, Vrijens B, Bosworth HB, Liu LZ, Zullig LL, Granger BB. Coming Full Circle in the Measurement of Medication Adherence: Opportunities and Implications for Health Care. *Patient Prefer Adherence* (2017). 11:1009–17. doi:10.2147/ppa.s127131
71. Dobbels F, Berben L, De Geest S, Drent G, Lennerling A, Whittaker C, et al. The Psychometric Properties and Practicability of Self-Report Instruments to Identify Medication Nonadherence in Adult Transplant Patients: A Systematic Review. *Transplantation* (2010). 90:205–18. doi:10.1097/tp.0b013e3181e346cd
72. The Basel Assessment of Adherence to Immunosuppressive Medication Scale (BAASIS®). (2022). Available at: http://baasis.nursing.unibas.ch/. (Accessed January 31, 2022)

# Transplant International

Official journal of the European Society for Organ Transplantation


## Editorial Office

Avenue du Tribunal Fédéral 34
CH – 1005 Lausanne
Switzerland

Tel +41 (0)21 510 17 40
Fax +41 (0)21 510 17 01

tieditorialoffice@frontierspartnerships.org
frontierspartnerships.org/journals/transplant-international

frontiers | Publishing Partnerships