Clinical Epidemiology

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

On the Importance of a Clear Definition of Time Horizon for Time-to-Event Dynamic Predictions: A Systematic Review and a Concrete Illustration in Kidney Transplantation

Lucas Chabeau[1,2], Vincent Bonnemains[1], Pierre Rinder[2], Magali Giral[3,4], Solène Desmée[5], Etienne Dantan[1]

[1]Nantes Université, Univ Tours, INSERM UMR 1246, MethodS in Patients-Centered Outcomes and HEalth Research, SPHERE, Nantes, France; [2]SEMEIA, Paris, Toulouse, France; [3]Nantes Université, CHU Nantes, INSERM, Center for Research in Transplantation and Translational Immunology, UMR 1064, ITUN, F-44000, Nantes, France; [4]Centre d'Investigation Clinique en Biothérapie, Nantes, France; [5]Univ Tours, Nantes Université, INSERM UMR 1246, MethodS in Patients-centered outcomes and HEalth Research, SPHERE, Tours, France

Correspondence: Etienne Dantan, Nantes Université, Univ Tours, INSERM UMR 1246, MethodS in Patients-centered outcomes and HEalth Research, SPHERE, IRS2, 22 Boulevard Bénoni Goullin, Nantes, 44200, France, Tel +33 2 53 00 91 28, Email Etienne.Dantan@univ-nantes.fr

Clinical Epidemiology downloaded from https://www.dovepress.com/
For personal use only.

Introduction: Dynamic predictions estimate survival probability until a given horizon, conditionally on being event-free at landmark times and additional information on predictive variables available at these times. The horizon may be defined as a *final time horizon* or the end of a *sliding horizon window*.

Methods: Following PRISMA, CHARMS, and TRIPOD recommendations, a systematic review of dynamic predictions, querying Medline in May 2025 with no date restriction and including only English articles, assessed heterogeneity in time horizon reporting. Moreover, from 2,523 kidney recipients, prognostic capacities of the Dynamic predictions of Patient and kidney Graft survival have been studied considering either a *final time horizon* or a *sliding horizon window*.

Results: From 171 articles included in the systematic review, 97 articles (57%) used a *sliding horizon window*, 36 (21%) a *final time horizon* and 38 articles (22%) had no clear time horizon definition. From the kidney transplant recipients' sample, discrimination and calibration of dynamic predictions are not comparable between the two time horizon definitions due to differing event incidence depending on the nature of the prediction window that can be either reduced or sliding. For a 5-year *sliding horizon window*, discrimination slightly increased with landmark times, and calibration appeared reasonable, especially at earliest landmark times. For an 11-year *final time horizon*, discrimination was high for earliest landmark times and increased over time, while calibration revealed underestimated predictions for earliest landmark times and overestimated for later ones.

Discussion: Our systematic review highlights the need for clearer time horizon reporting due to heterogeneous definitions. Our renal transplantation illustration revealed that the prognostic performances of Dynamic predictions of Patient and kidney Graft survival differ given the time horizon definition. More broadly, this study highlights the need for caution when interpreting prognostic performances, as it depends on whether the horizon window is sliding or reduced.

Keywords: time-to-event dynamic predictions, landmark times, horizon window, time horizon, discrimination, calibration

Introduction

In the personalized medicine era,[1] time-to-event dynamic predictions are becoming more widespread. They are defined as the probability of being event-free until a defined time (*time horizon*) given being event-free at the time of making a prediction (*landmark time*) and given available predictive variables at such prediction times.[2–5] Dynamic predictions take into account the valuable information consisting of the entire marker trajectory known at landmark time and have shown their importance in improving time-fixed predictions that only consider baseline variables.[5–7] The dynamic predictions are updated predictions whenever additional longitudinal data becomes available during the patient follow-

https://doi.org/10.2147/CLEP.S572750
Received: 9 October 2025
Accepted: 30 April 2026
Published: 14 May 2026

© 2026 Chabeau et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v4.0) License (http://creativecommons.org/licenses/by-nc/4.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

up. For example, Teramukai et al focused on dynamically predicting the cardiovascular endpoints of hypertensive patients until 3 years after inclusion using repeated on-treatment blood pressure measurements.[8] Ben-Hassen et al proposed dynamically predicting dementia based on longitudinal neurocognitive tests over the next 5 years following the time of making prediction.[9]

In such dynamic prediction context, the prediction window (*horizon window*) corresponds to the delay between the landmark time and the time horizon. The *horizon window* requires precisions in reporting since it can be defined in several ways.[6,10] One objective could be to predict the survival probability until a *final time horizon*, as, for instance, in Teramukai et al.[8] An alternative objective could be to predict the survival probability until the end of a *sliding horizon window*[6,11,12] as, for instance, in Ben-Hassen et al.[9] Whatever the objective and the chosen definition of the horizon window, prognostic scores require good prognostic performances to be useful.[13,14] They can be studied through global performances using time-dependent Brier Score or R^2-curve,[11,15] discrimination property using time-dependent AUC of ROC curves,[15,16] and calibration property using time-dependent calibration plots.[17]

The literature on dynamic predictions is growing; however, to the best of our knowledge, a clear definition of the prediction window corresponding to the two aforementioned objectives does not yet exist, and a thorough comparison between the two approaches has not been reported. Such topics are of crucial importance when assessing the prognostic performances, as the same metrics have different interpretations depending on the objective.

The objective was to illustrate that the prognostic performances and their interpretations differ given the two *time horizon* definitions. We presented the mathematical framework to obtain dynamic predictions according to the two definitions (Section 2). We then conducted a systematic review of articles concerning dynamic predictions to objectively assess how the horizon window was reported in the literature (Section 3). By reanalysing kidney recipient data previously used to develop and internally validate the Dynamic predictions of Patient and kidney Graft survival (DynPG),[17] we illustrated that the prognostic performances of dynamic predictions differ between time horizon definitions (Section 4). Finally, Section 5 offers a discussion and conclusions.

Methods

Notations

Let us assume a learning sample of n independent and identically distributed patients with the observed data $\{Y_i, t_i, T_i, \delta_i, A_i; i = 1, \ldots, n\}$. Here, $Y_i = \{Y_{ij}; j = 1, \ldots, n_i\}$ corresponds to the set of n_i longitudinal marker values measured at the corresponding times $t_i = \{t_{ij}; j = 1, \ldots, n_i\}$ in individual i. Let us consider $T_i = \min(T_i^*, C_i)$ with T_i^* the true time-to-event, C_i the censoring time for subject i, and $\delta_i = 1\{T_i^* \leq C_i\}$ the event indicator function taking 1 if the event is not censored and 0 otherwise. We note $A_i = \{A_{ik}; k = 1, \ldots, K\}$ the K baseline variables. From a learning sample of n patients, we can estimate a prediction model of parameters θ. Among those, landmarking and joint modeling of longitudinal and survival data are popular approaches.[2,18,19]

Dynamic Predictions Definitions

The dynamic prediction π_l for a new patient l is defined as the probability of being event-free until a *time horizon u* given being event-free at the landmark time s ($s < u$), given baseline variables A_l and the longitudinal marker history available at time s (ie $\tilde{Y}_l(s) = \{Y_l(s_1), \ldots, Y_l(s_p); 0 \leq s_1 < \ldots < s_p < s\}$):

$$\pi_l(s, u; \theta) = \Pr\left(T_l^* > u \middle| T_l^* > s, A_l, \tilde{Y}_l(s); \theta\right) \quad (1)$$

We report two different uses of this generic definition of dynamic predictions,[2,3,5,10,20] that depend on the definition of *time horizon u* and that would not have the same clinical objective.

Dynamic Predictions Given a Final Time Horizon

The objective would be to dynamically predict the survival probability π_l for a patient l until a fixed *final time horizon U* (Figure 1). It corresponds to the probability to not suffer the event on a horizon window defined as the delay between the landmark time s and the *final time horizon U*:

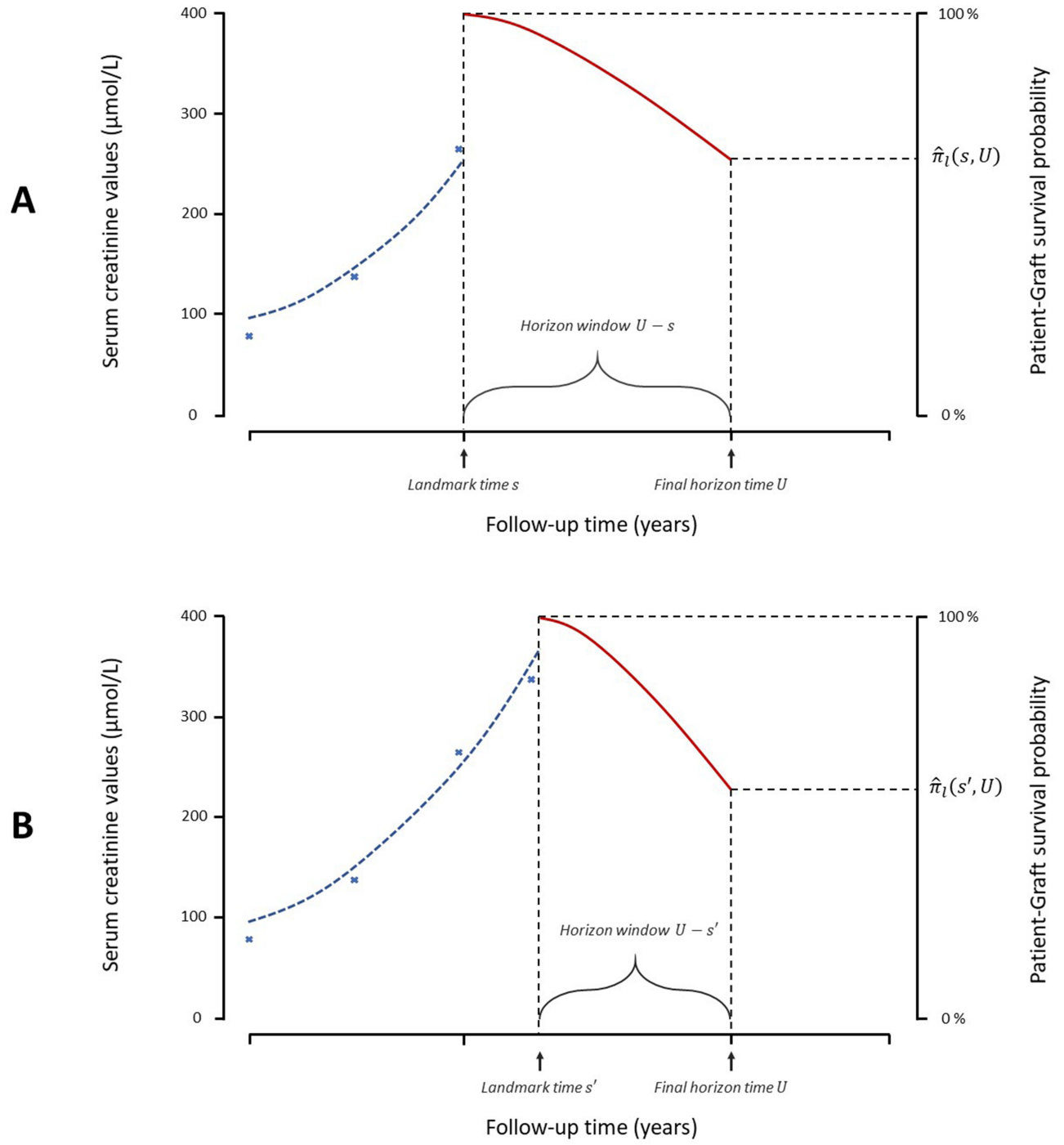

Figure 1 Scheme of dynamic predictions of patient and kidney graft survival until a final time horizon U (red line) given a landmark time s Part (**A**) and given a landmark time Part s' $(s' > s)$ (**B**), with longitudinal measures (blue crosses) and predicted marker evolution available before the landmark time (blue line).

$$\pi_l(s, U; \theta) = \Pr\left(T_l^* > U \middle| T_l^* > s, A_l, \tilde{Y}_l(s); \theta\right)$$

Using the estimated parameters $\hat{\theta}$ of the prediction model, the dynamic predictions considering a *final time horizon* can be estimated as the ratio between the survival probability until the *final time horizon* U and the survival probability until the time of making prediction s:

https://doi.org/10.2147/CLEP.S572750

$$\hat{\pi}_l(s,U;\hat{\theta}) = \frac{Pr(T_l^*>U|A_l,\tilde{Y}_l(s);\hat{\theta})}{Pr(T_l^*>s|A_l,\tilde{Y}_l(s);\hat{\theta})}$$

Following this, the horizon window $\{U-s\}$ is progressively reduced when landmark time s increases (Figure 1). We note that for a fixed $\tilde{Y}_l(s)$, the smaller the difference between s and U, the bigger $\hat{\pi}_l(s,U)$ will be. Indeed, $\pi_l(s,U;\theta)$ is bounded with:

$$\lim_{s\to 0}\hat{\pi}_l(s,U;\hat{\theta}) = \frac{Pr(T_l^*>U|A_l,\tilde{Y}_l(0);\hat{\theta})}{1} = Pr(T_l^*>U|A_l,\tilde{Y}_l(0);\hat{\theta})$$

$$\lim_{s\to U}\hat{\pi}_l(s,U;\hat{\theta}) = \frac{Pr(T_l^*>U|A_l,\tilde{Y}_l(U);\hat{\theta})}{Pr(T_l^*>U|A_l,\tilde{Y}_l(U);\hat{\theta})} = 1$$

Dynamic Predictions Given a Sliding Horizon Window

Some authors have proposed an alternative definition of the dynamic predictions.[6,11,12] The target of inference of a patient l is to predict the survival probability until the end of a *horizon window* of length Δt:

$$\pi_l(s,s+\Delta t;\theta) = \Pr(T_l^*>s+\Delta t \,|\, T_l^*>s, A_l, \tilde{Y}_l(s);\theta)$$

Following this definition, the *horizon window* Δt is thus sliding when updating the prediction, ie. the landmark time increases (Figure 2). Using the estimated parameters $\hat{\theta}$ of the prediction model, the dynamic predictions given a *sliding horizon window* can be decomposed as the ratio between the survival probability until the end of the *horizon window* $s+\Delta t$ and the survival probability until the time of making prediction s:

$$\hat{\pi}_l(s,s+\Delta t;\hat{\theta}) = \frac{Pr(T_l^*>s+\Delta t|A_l,\tilde{Y}_l(s);\hat{\theta})}{Pr(T_l^*>s|A_l,\tilde{Y}_l(s);\hat{\theta})}$$

Accuracy Measures for Dynamic Predictions

The global prognostic performance of dynamic predictions can be assessed by measuring the prediction error,[21] with the popular Brier Score metric, for instance. The time-dependent expected Brier Score is defined as follows:[15,18]

$$BS(s,u) = E\left[(1(T^*>u)-\pi(s,u))^2|T^*>s\right]$$

The Brier Score is a mean square error term. The closer the prediction is to the observation, the closer the Brier Score is to 0. When considering the *final time horizon*, it means that the survival probability until a *final time horizon* is expected to be close to the observed event indicator after the final *time horizon* $\left(BS(s,U) = E\left[(1(T^*>U)-\pi(s,U))^2|T^*>s\right]\right)$. When considering the *sliding horizon window*, it means that the survival probability after an *horizon window* Δt is expected to be close to the observed event indicator after this *horizon window* $\left(BS(s,s+\Delta t) = E\left[(1(T^*>s+\Delta t)-\pi(s,s+\Delta t))^2|T^*>s\right]\right)$. In a dynamic context, one limit of the Brier Score is that it is sensitive to the marginal event probability that could change as the landmark time increases. Van Houwelingen et al proposed studying the relative error reduction,[21] also named R^2-curve:[11]

$$R^2(s,u) = 1 - \frac{BS(s,u)}{BS_0(s,u)}$$

where $BS(s,u)$ is the Brier Score obtained using the model of interest and $BS_0(s,u)$ is the Brier Score of a reference model (Kaplan–Meier estimator for instance) allowing to assess the global performance of dynamic predictions whatever the evolution of the marginal event probability along the landmark times. Since the Brier Score computation would be impacted by the consideration of the *final time horizon* or the *sliding horizon window*, the R^2-curve would also not be the same between the two approaches.

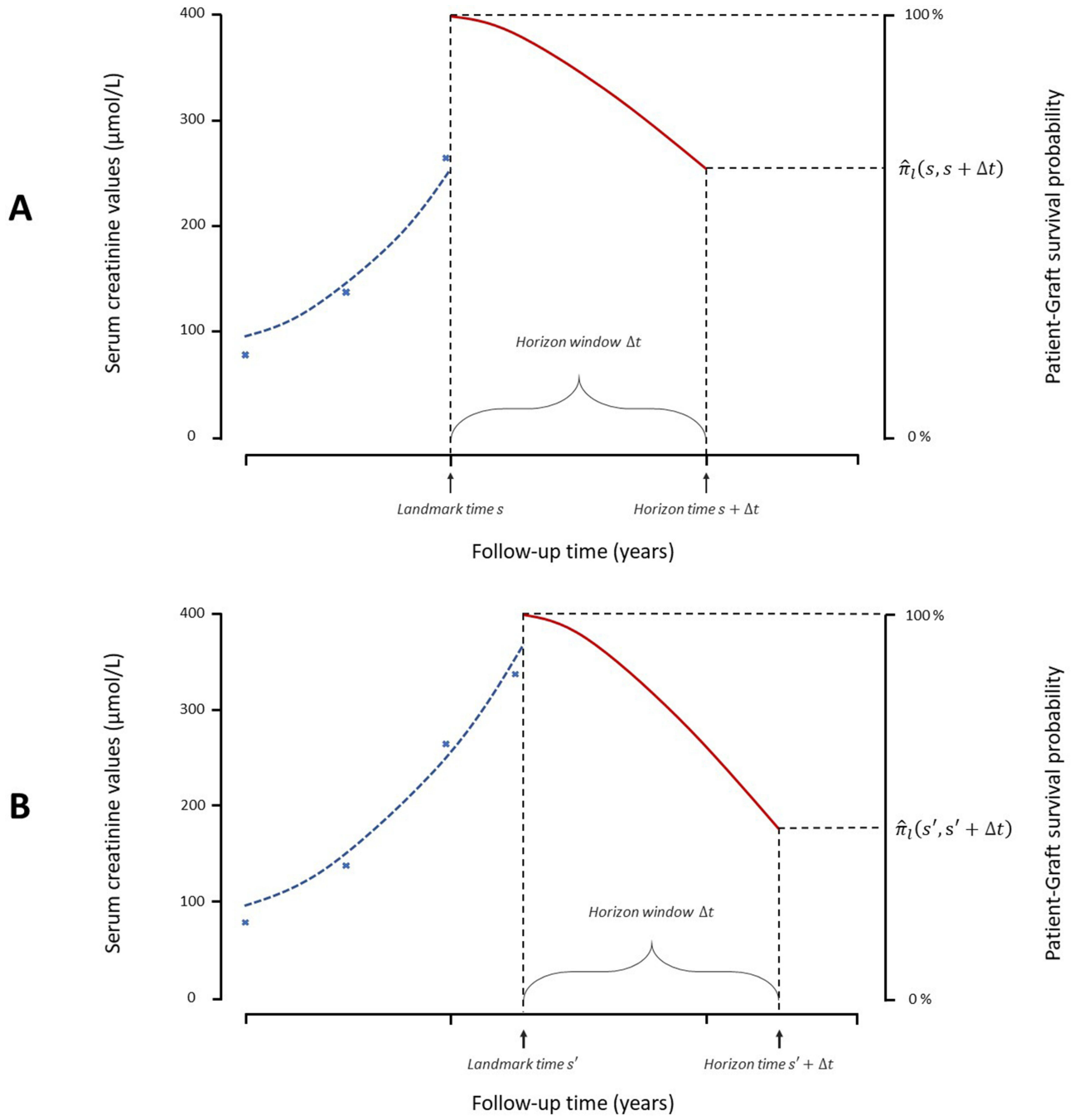

Figure 2 Scheme of dynamic predictions of patient and kidney graft survival for a fixed horizon window Δt (red line) given landmark time s Part (**A**) and given landmark time $s'(s'>s)$ Part (**B**), with longitudinal measures (blue crosses) and predicted marker evolution available before the landmark time (blue line).

Discrimination is the ability of a prognostic tool to order the risk between subjects. To assess discrimination, the well-known Area Under the Receiver Operating Characteristics Curve (AUC) can be explicitly defined in a dynamic context as:[15,16]

$$AUC(s,u) = \Pr\left(\pi_l(s,u) > \pi_{l'}(s,u) \middle| T_l^* > u, s < T_{l'}^* < u\right)$$

When considering the *final time horizon*, the AUC corresponds to the probability that a subject l at-risk at landmark time s and who would not suffer the event before the *final time horizon* U would have a higher survival prediction than a subject l' at-risk at landmark time s and who would suffer the event before the *final time horizon* U $\left(AUC(s,U) = \Pr\left(\pi_l(s,U) > \pi_{l'}(s,U) \middle| T_l^* > U, s < T_{l'}^* < U\right)\right)$. When considering the *sliding horizon window*, the AUC corresponds to the probability that a subject l at-risk at landmark time s and who would not suffer the event before the end of the

https://doi.org/10.2147/CLEP.S572750

horizon window $s + \Delta t$ would have a higher survival prediction than a subject l' at-risk at landmark time s and who would suffer the event between s and $s + \Delta t$ $\left(AUC(s, s + \Delta t) = \Pr\left(\pi_l(s, s + \Delta t) > \pi_{l'}(s, s + \Delta t) \middle| T_l^* > s + \Delta t, s < T_{l'}^* < s + \Delta t\right)\right)$.

The calibration property assesses the ability of a prognostic tool to provide a prediction close to the observed outcome. Usually assessed through calibration plots, the calibration is described by comparing the predicted values within subgroups (defined from quantiles of predictions) to the observed survival probabilities (computed using the Kaplan–Meier estimator for instance). When considering the *final time horizon*, a good calibration means that, for any given x value, we expect that among all subjects who are predicted to survive with a x % probability, x out of 100 will not experience the event before the *final time horizon*. When considering the *sliding horizon window*, a good calibration means that, for any given x value, we expect that among all subjects who are predicted to survive with a x % probability, x out of 100 will not experience the event before the end of the *horizon window*.

Systematic Review of Dynamic Predictions

Search Strategy

We conducted a systematic review following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement - PRISMA (Table S1).[22] We searched the Medline database on 5th May 2025, with no date restriction and including only English articles. The search equation used is reported in Appendix 1.

Study Selection

As this systematic review questioned the definition and the choice of time horizon, we included all articles focusing on the development or the validation of individual dynamic prediction scoring systems as well as methodological papers concerning dynamic predictions. Two reviewers (LC and ED) independently screened references by title and abstract, and their concordance was assessed based on the proportion of observed agreements and Cohen's kappa. Our exclusion criteria were reviews or meta-analyses, full texts not found or conference abstracts and editorials.

Data Extraction

Each article was randomly allocated to two of the five reviewers (LC – VB – PR – SD – ED). Following the CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies (CHARMS),[23] we predefined a standardized form for data extraction. Whether it was a methodological article or an article concerning the development or the validation of dynamic predictions, for each eligible article, we collected descriptive information as the author's name, year and journal, the data source. If we finally decided to not include it, the reason for was specified. All eligible texts were independently assessed by the two reviewers for inclusion/exclusion, and for time horizon classification when included. The inter-rater agreement between reviewers was measured based on the proportion of observed agreement and Cohen's kappa. For each included article, the time horizon was qualified as either a *final time horizon* or the end of a *sliding horizon window* based on an explicit equation, explicit texts, and/or a clear graphic representing individual dynamic prediction, or qualified as *unclear* information when such explicit information was unavailable or contradictory information was present. When necessary, any discrepancies were resolved through discussion with another reviewer to reach a consensus.

We also collected the study design, the sample size, the population characteristics, the predicted outcomes, the prediction times (ie. landmark times corresponding to the times when one calculates the prediction), the time horizon (ie. the end of the prediction time window), the predictive tools and the development details (such as the methodology used, the variables of the scoring systems). The TRIPOD statement recommends to develop a prognostic score from a learning sample and to validate the prognostic performances from an independent internal and/or external validation sample to avoid reporting the prognostic performances on the learning sample only, which would lead to overestimating the performances.[24,25] We thus extracted information on internal and external validations as well.

Results

The search identified 425 unique articles. Screening of titles and abstracts identified 275 papers eligible for full-text review. As detailed on the flow diagram (Figure 3), 104 articles were excluded. Details of the inter-rater reliability are

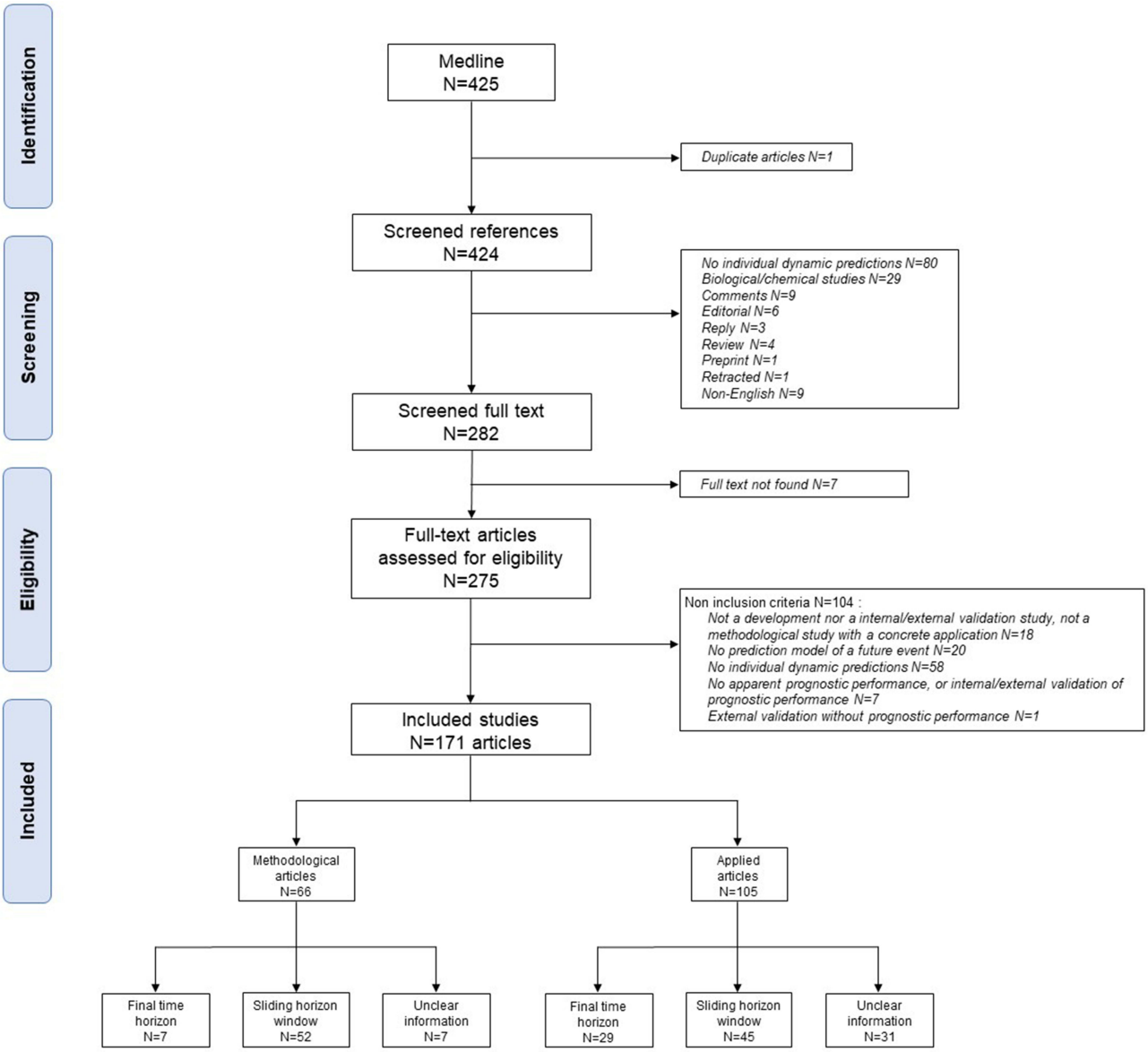

Figure 3 PRISMA flow diagram, selection of included studies in the systematic review.

provided in Table S2. Finally, we included a total of 171 articles (Appendix 2), among which 66 (39%) were methodological articles and 105 (61%) were applied articles.

A general overview of the included articles is presented. Dynamic predictions appear to be a relatively recent research topic in biostatistics with 169 (99%) articles published after 2010. Sixty-six (39%) of the retained articles are methodological studies that proposed new modeling approach. This systematic review confirmed the dominance of landmarking (n=69, 40%) and joint modeling for longitudinal and survival data (n=66, 39%) to develop dynamic predictions. To a lesser extent, we identified the emergence of machine learning approaches (n=33, 19%), with, for instance, random survival forest,[26–28] or neural networks.[29,30] All articles referring to machine learning approaches were published after 2018. This will probably increase in the future with the fast development of such modeling approaches. Among the applied articles (n=105, 61%), we also observed that 22 studies (21%) only presented a development modeling of dynamic prediction without validation of their predictive performance on an independent dataset (Table 1). This does not respect the TRIPOD recommendations about the right process to develop and validate prognostic

https://doi.org/10.2147/CLEP.S572750

Table 1 Description of the Articles Included in the Systematic Review

	All Included Articles N=171	Methodological Articles N=66	Applied Articles N=105
Time prediction horizon definition			
Final time horizon	36 (21%)	7 (11%)	29 (28%)
Sliding horizon window	97 (57%)	52 (79%)	45 (43%)
Unclear	38 (22%)	7 (11%)	31 (30%)
Published after 2010	169 (99%)	66 (100%)	103 (98%)
Main statistical analyses for model development			
Joint model	66 (39%)	36 (55%)	30 (29%)
Landmarking	69 (40%)	22 (33%)	47 (46%)
Machine Learning	33 (19%)	7 (11%)	26 (25%)
Others	16 (9%)	7 (11%)	9 (9%)
Global performance	76 (44%)	45 (78%)	31 (30%)
Brier Score	50 (66%[a])	33 (73%[a])	17 (55%[a])
Prediction error	29 (38%[a])	15 (33%[a])	14 (45%[a])
Others	4 (5%[a])	3 (7%[a])	1 (3%[a])
Discrimination	148 (87%)	47 (81%)	101 (99%)
AUC	124 (84%[a])	43 (91%[a])	81 (80%[a])
C-index	27 (18%[a])	6 (13%[a])	21 (21%[a])
Others	10 (7%[a])	1 (2%[a])	9 (9%[a])
Calibration	51 (30%)	8 (14%)	43 (42%)
Calibration plot	40 (78%[a])	7 (88%[a])	33 (77%[a])
Calibration slope	3 (6%[a])	0 (0%[a])	3 (7%[a])
Heuristic shrinkage factor	6 (12%[a])	0 (0%[a])	6 (14%[a])
Others	7 (14%[a])	1 (13%[a])	6 (14%[a])
Type of study			
Development alone	-	-	22 (21%)
Development & Internal validation	-	-	58 (55%)
Development & External validation	-	-	8 (8%)
Development, Internal & External validation	-	-	15 (14%)
External validation alone	-	-	2 (2%)

Note: [a]Percentage among the count of articles reporting the prognostic ability.

performances. Nevertheless, 58 articles included internal validation, 8 articles included an external validation and 15 both types of validation, leading to 66% of the applied articles in agreement with the TRIPOD.

According to the main objective of our systematic review, the different approaches used to define prediction horizons were examined. Among the included articles, we identified 36 (21%) articles that defined dynamic predictions using a *final time horizon*, 97 (57%) articles using *sliding horizon windows*, and 38 (22%) articles in which the definition of time horizon was unclear (Table 1). We observed that this distribution differed between methodological and applied articles. Methodological articles mainly used *sliding horizon windows* for dynamic predictions (n=52, 79%), while only 7 (11%) articles used a *final time horizon*, and 7 (11%) articles did not clearly define the approach adopted. Regarding applied articles, the definitions used were much more heterogeneous, as 45 (43%) used *sliding horizon windows*, 29 (28%) used a *final time horizon*, and 31 (30%) were unclear regarding the approach used.

Beyond the description of the time horizon considered, our systematic review allows to appraise the reporting quality on dynamic predictions. As clearly recommended in the literature,[13] assessing prognostic performances is of major importance for prediction tools to propose useful scores. Discrimination capacity is well reported (n=148, 87%) with AUC as the principal performance indicator. Global performances (n=76, 44%) are less often provided, with Brier Score the most

frequent indicator. Calibration (n=51, 30%) is also less reported, with calibration plot the most used approach. Details of such prognostic performances are also not homogeneous between methodological and applied articles (Table 1).

Application

Context

In kidney transplantation, patients are particularly interested in their kidney graft survival, before any risk of adverse outcomes or infections.[31] In this context, we developed and internally and externally validated "Dynamic predictions of Patient and kidney Graft survival" (DynPG) for kidney recipients alive with a functioning graft at 1-year post-transplantation.[17,32] The main outcome was the delay from 1-year post-transplantation to patient and kidney graft failure defined as the first event between return-to-dialysis, pre-emptive retransplantation and death with a functioning graft. Such dynamic predictions can be obtained from six baseline variables: recipient age, graft rank, cardiovascular histories, pretransplantation anti-HLA class I immunization, serum creatinine at 3-months post-transplantation, occurrence of acute rejection in the first year post-transplantation, and also the complete longitudinal serum creatinine trajectory available at the time of prediction. Our aim was to provide patients and physicians with information regarding mid-term prognosis by considering a *sliding horizon window* of 5 years. In line with this clinical objective, the DynPG was defined as the probability of being kidney graft failure-free over the 5 years following the landmark time, for each landmark time from 1 to 6 years post-transplantation. We retained 6 years post-transplantation as the maximum landmark time because there was still a reasonable number of at-risk patients at the end of the prediction window (178 patients still at risk of patient and kidney graft failure at 11 = 6 + 5 years post-transplantation in the validation sample).

We aimed to illustrate that the prognostic performances would not be the same under the assumptions of a *sliding horizon window* or a *final time horizon*. While a *sliding horizon window* may be relevant to monitor the patient risk along his/her follow-up with always the same *horizon window* of 5 years, the clinical objective is different under the prism of a *final time horizon*. Regarding the *final time horizon*, we aim to provide long-term dynamic predictions with good confidence up to the *final time horizon*. Therefore, we chose the *final time horizon* as the maximum time considered in the *sliding horizon window* (ie. 11 years). Benefiting from the previously estimated joint model and the same internal validation sample, we compared the prognostic performances of the DynPG under the assumption of a 5-year *sliding horizon window* and under the assumption of an 11-year *final time horizon*.

Study Population

Data were extracted from the French multicentric observational and prospective DIVAT cohort (Données Informatisées et VAlidées en Transplantation; (www.divat.fr). The study was conducted in accordance with the Declaration of Helsinki and approved by the institutional review board of the DIVAT network. Data confidentiality was ensured based on the recommendations of the French Commission for Data Protection (Commission Nationale Informatique et Liberté, CNIL no. 914184, ClinicalTrials.gov recording NCT02900040). All participants provided written informed consent.

The inclusion criteria were adult recipients who received a first or second renal graft, transplanted between January 2000 and October 2016, from a living or heart-beating deceased donor, who were alive with a functioning graft at 1 year post-transplantation and maintained under tacrolimus and mycophenolate. The extracted DIVAT cohort data consisted of a learning set of 2,749 patients, initially used to estimate the shared-random joint model, and an internal validation set of 2,589 patients. Due to missing data for variables required to calculate the predictions, 66 patients (less than 3% of the eligible patients) were excluded from the analysis of the validation set (n = 2,523). However, the excluded patients had characteristics comparable to those of the included patients, suggesting that the missing data process was random; more details on this validation sample are reported in Fournier et al.[17]

Results

Under the 5-year *sliding horizon window* setting, the results were as previously published.[17] The global prognostic performance of the DynPG appeared relatively stable along the landmark times with R^2 values ranging from 14% (95% CI 7–21%) to 15% (95% CI −2% to 33%) at 1 and 6 years post-transplantation, respectively (Figure 4A). The

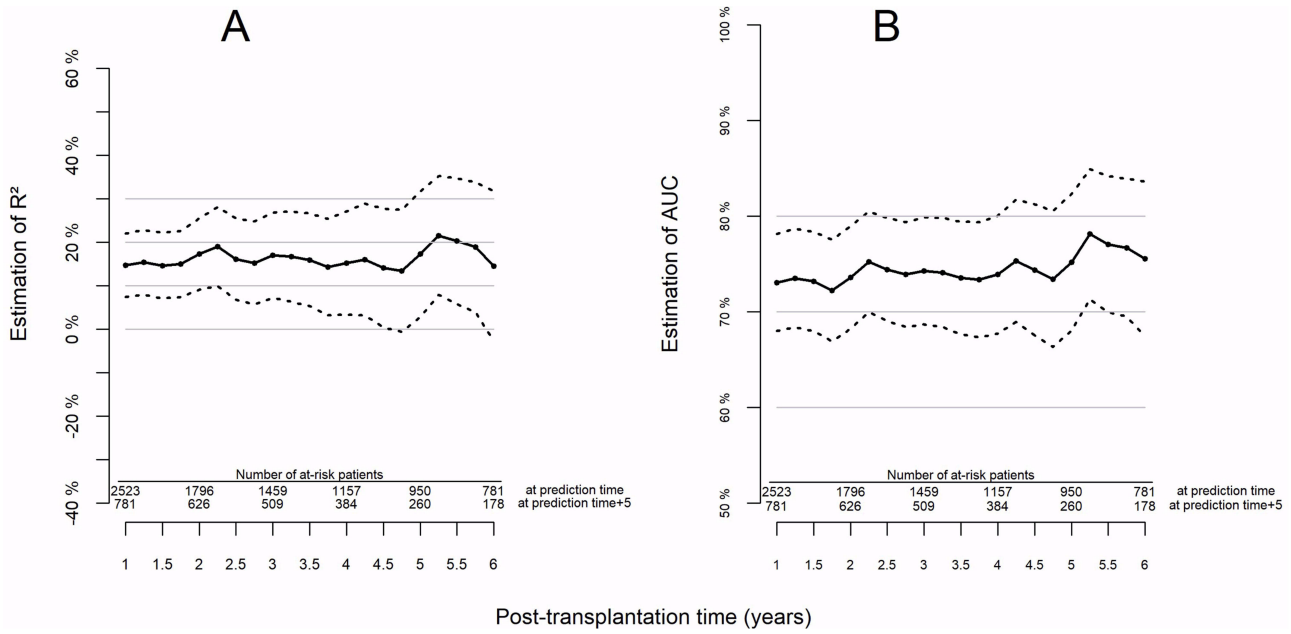

Figure 4 Prognostic capacities of the dynamic predictions obtained from the DIVAT internal validation sample (n=2,523, 66 observations deleted due to missing data concerning covariates) for landmark times varying from 1 to 6 years post-transplantation for a given 5-year *horizon window;* R^2 evaluated global performance (**A**) and the AUC appraised the discrimination accuracy (**B**). Estimations are drawn as solid lines and the corresponding 95% CIs are drawn as dashed lines.

discrimination slightly increased along the prediction times with the AUC values ranging from 0.72 (95% CI 0.67–0.78) to 0.76 (95% CI 0.68–0.85) at 1 and 6 years post-transplantation (Figure 4B). The calibration properties appeared reasonable (Figure S1). Note that, because the discrimination performances increased and the global prognostic performances remained stable, the calibration properties decreased for the late landmark times. As the landmark times increased, the calibration slope progressively deteriorated, with the estimated intercept and slope moving further away from 0 and 1, respectively (Figure S1).

Under the 11 years *final time horizon* assumption, the global prognostic performances of the DynPG were at a higher level for the earliest landmark times, but decreased along the landmark times. We estimated R^2 values of 22% (95% CI 12–32%), 15% (95% CI −2% to 31%) and −2% (95% CI −41% to 37%) at 1, 6 and 10 years post-transplantation, respectively (Figure 5A). The discrimination was high for the earliest landmark times and increased over times with AUC values of 0.76 (95% CI 0.70–0.82), 0.76 (95% CI 0.67–0.84) and 0.88 (95% CI 0.79–0.98) at 1, 6 and 10 years post-transplantation, respectively (Figure 5B). In contrast, we observed a very poor calibration plot, where the estimated intercept and slope of the calibration slope were far from 0 and 1, respectively, since the earliest landmark times (Figure S2). The predictions appeared underestimated for the earliest landmark times and overestimated for the subsequent ones.

Following these two definitions, the *horizon window* is either reduced or sliding as the landmark time increases. Therefore, the at-risk population at landmark times are identical given these two definitions. However, the number of observed events and the number of censored subjects in each window differ between the two approaches due to different horizon windows, resulting in a number of at-risk patients at the end of the horizon window that are not comparable (Figure S3, S4). Therefore, the incidence rates in the prediction window are not the same between the two approaches. Consequently, dynamic predictions cannot be compared, and they tell different stories. For instance, under the *sliding horizon window* assumption, at 1 year post-transplantation, we estimated a 72% probability that the predicted survival of a patient who actually experienced a graft failure within the 5 years was lower than that of a patient who did not (Figure 4 – part B). Under the *final time horizon* assumption, at 1 year post-transplantation, we estimated a 76% probability that the predicted survival of a patient who actually experienced a graft failure before 11 years post-transplantation was lower than that of a patient who did not (Figure 5 – part B). In terms of calibration properties, we may reasonably accept that DynPG was sufficiently well calibrated for the earliest landmark times under the 5-year

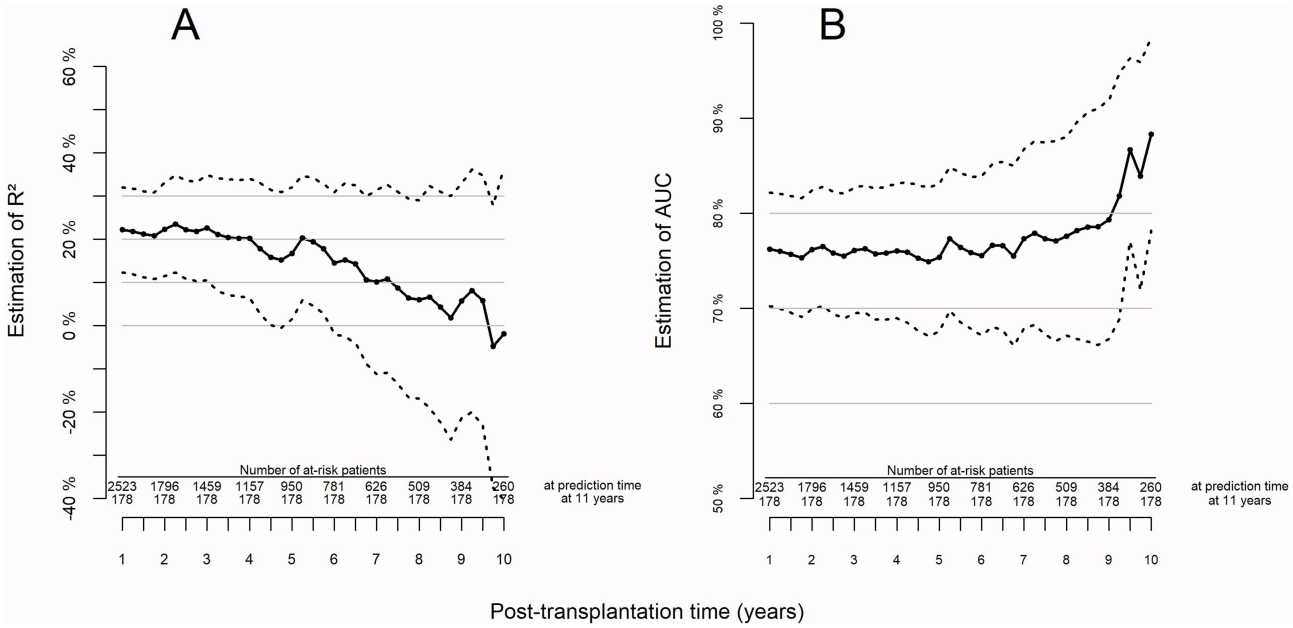

Figure 5 Prognostic capacities of the dynamic predictions obtained from the DIVAT internal validation sample (n=2,523, 66 observations deleted due to missing data concerning covariates) for landmark times varying from 1 to 10 years post-transplantation for a *final time horizon* of 11 years post-transplantation; R2 evaluated global performance (**A**) and the AUC appraised the discrimination accuracy (**B**). Estimations are drawn as solid lines and the corresponding 95% CIs are drawn as dashed lines.

sliding horizon window assumption (Figure S1). In contrast, under the 11-year *final time horizon* assumption, the calibration is quite poor (Figure S2).

Discussion

In this study, we distinguished two types of time horizon – *final time horizon* or end of a *sliding horizon window* – for dynamic predictions. We conducted a systematic review that stated the heterogeneity of the used time prediction horizons in the literature about dynamic predictions. While the two definitions are similar, a specific definition is of major importance since the prognostic performances obtained are different given the nature of the prediction window. This is also well illustrated by our concrete application in kidney transplantation.

The concept of P4-medicine (predictive, preventive, personalized and participatory) is now largely developed in the literature, but still difficult to apply in clinical practice.[1] Van Calster et al recently highlighted several challenges that make difficult the development, the validation and the implementation of prediction models for a clinical use[33] Dynamic predictive tools can be beneficial for such a health policy and promote shared medical decision making, provided that the prognostic performances are sufficiently good. While dynamic predictions are updated predictions whenever additional information is available during the patient follow-up,[6,10] their associated prognostic performances depend on the at-risk population at the time of making a prediction and can evolve as the landmark time increases. In this work, we insist on the fact that prognostic performances are also related to the prediction window. We showed that the incidence rates of the event would not be identical between the two time horizon definitions due to the nature of the window that can be either reduced or sliding, resulting in prognostic performances that are not comparable. The corresponding interpretations of prognostic performances should therefore be formulated with caution since they do not tell the same story in the two contexts.

In our concrete application to kidney transplantation, the dynamic predictions obtained with the *sliding horizon window* framework provide good calibration properties from 1-year until 6-year landmark times. In our opinion, these properties make the DynPG suitable for following and monitoring patient health evolution across landmark times and make it a promising option for individualized and personalized medicine.[17,32] The international SONG-Tx initiative highlights the importance of core outcome domains in kidney transplantation, as graft health and mortality remain among the most important criteria in the eyes of patients.[34] Therefore, the mid-term prognostic information provided by the DynPG could constitute individual levers for action. This effect could increase patient adherence to their treatment and increase patient

https://doi.org/10.2147/CLEP.S572750

empowerment, as patients can take an active role in their chronic disease management.[35,36] This could also help patients to better manage their feelings of uncertainty about the survival of their transplant in the not-too-distant future. Furthermore, we may envisage guiding patients through the care organization based on stratified survival probabilities.

When considering the *final time horizon* (ie. reduced window), the comparison of prognostic performances across the landmark times do not make sense since the lengths of the prediction window are not the same. Despite the bounded character of dynamic predictions for a *final time horizon*, it is important to note that such predictions are not always monotonic along the landmark times because of the actualization of the longitudinal marker. The amelioration or deterioration of the prognosis between two landmark times is not necessarily due to a difference in health state characterized by the new marker measurement. This is noised by the mathematical artefact brought by the reduction of the prediction window. Since it is difficult to know to what extent the prognosis evolution is due to the reduction of the window rather than to the marker evolution, a comparison cannot be done between predictions realized at two different landmark times, contrary to the *sliding horizon window* framework. For this reason, the *final time horizon* approach does not seem suited to follow the patient health evolution through survival probabilities and cannot be envisaged to support patient personalized and individualized care. In our kidney transplantation application, the poor calibration property across the landmark times for a *final time horizon* of 11 years post-transplantation do not allow to consider that individual survival probabilities are correctly predicted. The more the prediction time approaches 11 years, the more we predicted patient and kidney survival probabilities close to 1, which is not confirmed by the observed event repartition. This inevitably results in poor calibration. Based on the predicted patient and kidney graft survival probabilities, such an approach cannot be used to individualized the kidney recipient taking care along the follow-up. Nevertheless, in a context where the objective is not the individualisation of patient care, but rather the stratification of the studied population, considering a *final time horizon* may be of interest. In kidney transplantation, the TELEGRAFT study aimed to assess interest in telemedicine depending on the strata of patient graft failure risk calculated at baseline.[37] By extension, the satisfying levels of discrimination performance of the Dynamic predictions of Patient and kidney Graft survival can support stratified medicine. For example, healthcare visit schedules could be optimized by reducing monitoring or considering telemedicine consultations for low-risk patients while reinforcing follow-up for patients at higher risk.[37,38] From a public health perspective, such an efficient allocation of resources could also provide economic benefits. In the context of a chronic disease such as kidney transplantation, the definition of a clinically relevant final time horizon may be difficult from a patient-centred perspective. Other clinical contexts may be of interest when a *final time horizon* is already known. For example, manufacturers of medical devices are required to specify a maximum duration of use. The aortic valve duration or hip replacement duration could be predicted for such final time horizons announced by suppliers[39] For pregnancy outcomes, the prediction of adverse neonatal outcomes could be provided for a final time horizon of nine months.[40,41]

Considering the methodological differences between the two time horizon definitions, the studies' clinical objectives should be clearly anticipated to correctly assess the prognostic performances. Indirectly, the heterogeneity that we observed in the reporting of time horizon windows in our systematic review, which was more substantial in applied articles, reveals a lack of clarity regarding the clinical objectives justifying dynamic prediction development. Van Calster clearly stated poor or ambiguous reporting as one of the important limits for the use of prediction models.[33] They highly recommend to follow adequate reporting guidelines as TRIPOD,[24] or the recent TRIPOD+AI for which specific items concern the AI predictive algorithms.[42] Our systematic review also informed about the methodological quality of dynamic predictions. A non negligible proportion of the literature did not sufficiently respect the TRIPOD recommendation,[24] with a lack of validation studies for instance. We also noted that discrimination properties were mainly reported and referred to the stratification of the studied population, but calibration metrics that are essential for personalized care were not reported in numerous studies.[43] Such findings are consistent with previous literature demonstrating the lack of validation studies and the poor reporting of calibration performances.[44–46]

Our study has several limitations. First, we conducted a systematic review that suffered from several drawbacks. We performed it from a unique database (Medline). We did not use a formal risk-of-bias tool such as the Prediction model Risk Of Bias Assessment Tool (PROBAST), which is based on TRIPOD recommendations.[47,48] Instead, in line with TRIPOD, we developed a standardized data extraction form capturing information similar to the PROBAST items. We do

not believe this choice affected our conclusions, as the main objective of our systematic review was not to perform a meta-analysis, but rather to provide an overview of the types of horizon windows for dynamic prediction tools. Therefore, the main limitation of our systematic review is related to a possible publication bias issue, to which all systematic reviews are subject. Indeed, we cannot exclude the underrepresentation of external validation studies with non-confirmatory prognostic performances.[49] Second, the results of our illustration depend on the chosen prediction horizons. Fournier et al previously chose a 5-year sliding horizon window to provide mid-term prognosis for patients and physicians.[17] In this work, we adopted a similar choice for the sliding horizon window framework, while choosing an 11-year horizon for the final time horizon framework. We fully recognise that other choices could have been considered in accordance with other clinical interests. For example, a shorter horizon window of 1 year could help identify patients at particularly high risk and support earlier re-registration on waiting lists for pre-emptive re-transplantation. Nevertheless, we do not believe that this choice calls our conclusions into question. Indeed, other prediction window lengths would all the same shows that prognostic performances depend both on the population at risk at the time of prediction and on the event incidence within the prediction window, which by definition will vary depending on its nature (sliding or reduced). In practice, the nature of the prediction window and its horizon should be guided by clinical objectives. Third, the mathematical notations introduced and the following concrete application were restricted to the specific case of one longitudinal marker (eg. serum creatinine) and one clinical event (eg. patient and kidney graft survival), while more complex situations exist. Regardless, notations can be easily adapted, and our conclusions regarding time prediction windows remain valid for broader modelling frameworks. For example, numerous developments are related to improvements in the joint modelling of longitudinal and survival processes, contributing to the improvement of dynamic prediction performances. Considering competing risks,[50,51] refining the longitudinal evolution with flexible modeling,[52,53] or incorporating multiple longitudinal markers,[50,54] could enhance the prediction of clinical events of interest. For kidney transplantation, Chabeau et al recently proposed a joint model of the evolution of serum creatinine and donor-specific antibody immunisation, as well as their associations with the cause-specific risks of graft failure and death with a functioning graft,[55] which could be used to improve the DynPG.

Conclusions

In conclusion, the choice of the time prediction window was found to be crucial in dynamic prognostic context and should depend on clinical objectives. To the best of our knowledge, this study represents the first systematic review addressing the topic of time horizon definition and highlights the need for clearer reporting of time horizons in dynamic prediction studies, where definitions currently vary across the literature. Using kidney transplantation as an illustration, we demonstrated that the prognostic performances of Dynamic predictions of Patient and kidney Graft survival vary depending on how the time horizon is defined. More broadly, our findings emphasize the need for caution when interpreting prognostic performance, as it may differ according to whether the horizon window is defined as sliding or reduced.

Abbreviations

95% CI, 95% Confidence Interval; AUC, Area Under the Curve; CHARMS, CHecklist for critical Appraisal and data extraction for systematic Reviews of prediction Modelling Studies; DIVAT, Données Informatisées et VAlidées en Transplantation; DynPG, Dynamic predictions of Patient and kidney Graft survival; HLA, Human Leukocyte Antigen, PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses; ROC, Receiver Operating Characteristic; TRIPOD, Transparent Reporting of multivariable prediction model for Individual Prognosis Or Diagnosis statements.

Data Sharing Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Ethics Approval Statement and Informed Consent

Data confidentiality was ensured following the recommendations of the French commission for data protection (Commission Nationale Informatique et Liberté, CNIL no. 914184, ClinicalTrials.gov recording NCT02900040). The

present study was approved by the DIVAT scientific committee (Données Informatisées et VAlidées en Transplantation; www.divat.fr). Informed consent was obtained from all individual participants included in the study.

Acknowledgments

We thank members of the DIVAT consortium* and the clinical research assistant team (S. Le Floch, A. Petit, J. Posson, C. Scellier, V. Eschbach, K. Zurbonsen, C. Dagot, F. M'Raiagh, V. Godel, X. Longy, P. Przednowed). We are also grateful to Roche Pharma, Novartis, Sanofi and Astellas laboratories for supporting the DIVAT cohort (www.divat.fr) as the CENTAURE foundation (www.fondation-centaure.org).

Données Informatisées et VAlidées en Transplantation, DIVAT Cohort Collaborators (Medical Doctors, Surgeons, HLA Biologists)

Nantes: *Gilles Blancho, Julien Branchereau, Diego Cantarovich, Agnès Chapelet, Jacques Dantal, Clément Deltombe, Lucile Figueres, Claire Garandeau, Magali Giral, Caroline Gourraud-Vercel, Maryvonne Hourmant, Georges Karam, Clarisse Kerleau, Aurélie Meurette, Simon Ville, Christine Kandell, Anne Moreau, Karine Renaudin, Anne Cesbron, Florent Delbos, Alexandre Walencik, Anne Devis*; Paris-Necker: *Lucile Amrouche, Dany Anglicheau, Olivier Aubert, Lynda Bererhi, Christophe Legendre, Alexandre Loupy, Frank Martinez, Rébecca Sberro-Soussan, Anne Scemla, Claire Tinel, Julien Zuber*; Nancy: *Pascal Eschwege, Luc Frimat, Sophie Girerd, Jacques Hubert, Marc Ladriere, Emmanuelle Laurain, Louis Leblanc, Pierre Lecoanet, Jean-Louis Lemelle*; Lyon E. Hériot: *Lionel Badet, Maria Brunet, Fanny Buron, Rémi Cahen, Sameh Daoud, Coralie Fournie, Arnaud Grégoire, Alice Koenig, Charlène Lévi, Emmanuel Morelon, Claire Pouteil-Noble, Thomas Rimmelé, Olivier Thaunat*; Montpellier: *Sylvie Delmas, Valérie Garrigue, Moglie Le Quintrec, Vincent Pernin, Jean-Emmanuel Serre*

Funding

This work was partly supported by the French National Research Agency [grant number ANR PRC 2022 JM-QALYs]. Lucas Chabeau has received research support from SÊMEIA.

Disclosure

This paper has been uploaded to ResearchSquare as a preprint: https://www.researchsquare.com/article/rs-3938204/v1. Lucas Chabeau and Pierre Rinder report financial support as employees of Semeia. Other authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

1. Flores M, Glusman G, Brogaard K, Price ND, Hood L. P4 medicine: how systems medicine will transform the healthcare sector and society. *Pers Med*. 2013;10(6):565–576. doi:10.2217/pme.13.57
2. Rizopoulos D. *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*. CRC Press; 2012.
3. Barrett J, Su L. Dynamic predictions using flexible joint models of longitudinal and time-to-event data. *Stat Med*. 2017;36(9):1447–1460. doi:10.1002/sim.7209
4. Rizopoulos D, Taylor JMG. Optimizing dynamic predictions from joint models using super learning. *Stat Med*. 2024;43(7):1315–1328. doi:10.1002/sim.10010
5. Andrinopoulou ER, Harhay MO, Ratcliffe SJ, Rizopoulos D. Reflection on modern methods: dynamic prediction using joint models of longitudinal and time-to-event data. *Int J Epidemiol*. 2021;50(5):1731–1743. doi:10.1093/ije/dyab047
6. Proust-Lima C, Blanche P. Dynamic Predictions. In: *Wiley StatsRef: Statistics Reference Online*. John Wiley & Sons, Ltd; 2016:1–6. doi:10.1002/9781118445112.stat07876
7. Asgari S, Khalili D, Zayeri F, Azizi F, Hadaegh F. Dynamic prediction models improved the risk classification of type 2 diabetes compared with classical static models. *J Clin Epidemiol*. 2021;140:33–43. doi:10.1016/j.jclinepi.2021.08.026
8. Teramukai S, Okuda Y, Miyazaki S, Kawamori R, Shirayama M, Teramoto T. Dynamic prediction model and risk assessment chart for cardiovascular disease based on on-treatment blood pressure and baseline risk factors. *Hypertens Res off J Jpn Soc Hypertens*. 2016;39(2):113–118. doi:10.1038/hr.2015.120
9. Ben-Hassen C, Helmer C, Berr C, Jacqmin-Gadda H. Five-year dynamic prediction of dementia using repeated measures of cognitive tests and a dependency scale. *Am J Epidemiol*. 2022;191(3):453–464. doi:10.1093/aje/kwab269
10. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics*. 2011;67(3):819–829. doi:10.1111/j.1541-0420.2010.01546.x
11. Fournier MC, Dantan E, Blanche P. An R2 -curve for evaluating the accuracy of dynamic predictions. *Stat Med*. 2018;37(7):1125–1133. doi:10.1002/sim.7571

12. Asar Ö, Fournier MC, Dantan E. Dynamic predictions of kidney graft survival in the presence of longitudinal outliers. *Stat Methods Med Res*. 2021;30(1):185–203. doi:10.1177/0962280220945352
13. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. *Epidemiol Camb Mass*. 2010;21(1):128–138. doi:10.1097/EDE.0b013e3181c30fb2
14. Collins GS, Dhiman P, Ma J, et al. Evaluation of clinical prediction models (part 1): from development to external validation. *BMJ*. 2024;384:e074819. doi:10.1136/bmj-2023-074819
15. Blanche P, Proust-Lima C, Loubère L, Berr C, Dartigues JF, Jacqmin-Gadda H. Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics*. 2015;71(1):102–113. doi:10.1111/biom.12232
16. Kamarudin AN, Cox T, Kolamunnage-Dona R. Time-dependent ROC curve analysis in medical research: current methods and applications. *BMC Med Res Methodol*. 2017;17(1):53. doi:10.1186/s12874-017-0332-6
17. Fournier MC, Foucher Y, Blanche P, et al. Dynamic predictions of long-term kidney graft failure: an information tool promoting patient-centred care. *Nephrol Dial Transplant off Publ Eur Dial Transpl Assoc - Eur Ren Assoc*. 2019;34(11):1961–1969. doi:10.1093/ndt/gfz027
18. Ferrer L, Putter H, Proust-Lima C. Individual dynamic predictions using landmarking and joint modelling: validation of estimators and robustness assessment. *Stat Methods Med Res*. 2019;28(12):3649–3666. doi:10.1177/0962280218811837
19. Goldstein BA, Pomann GM, Winkelmayer WC, Pencina MJ. A comparison of risk prediction methods using repeated observations: an application to electronic health records for hemodialysis. *Stat Med*. 2017;36(17):2750–2763. doi:10.1002/sim.7308
20. Li K, Luo S. Dynamic predictions in Bayesian functional joint models for longitudinal and time-to-event data: an application to Alzheimer's disease. *Stat Methods Med Res*. 2019;28(2):327–342. doi:10.1177/0962280217722177
21. van Houwelingen HC, Putter H. *Dynamic Prediction in Clinical Survival Analysis*. CRC Press; 2012.
22. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med*. 2009;151(4):264–269,W64. doi:10.7326/0003-4819-151-4-200908180-00135
23. Moons KGM, de Groot JAH, Bouwmeester W, et al. Critical appraisal and data extraction for systematic reviews of prediction modelling studies: the CHARMS checklist. *PLoS Med*. 2014;11(10):e1001744. doi:10.1371/journal.pmed.1001744
24. Moons KGM, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162(1):W1–73. doi:10.7326/M14-0698
25. Steyerberg EW, Bleeker SE, Moll HA, Grobbee DE, Moons KGM. Internal and external validation of predictive models: a simulation study of bias and precision in small samples. *J Clin Epidemiol*. 2003;56(5):441–447. doi:10.1016/S0895-4356(03)00047-7
26. Devaux A, Genuer R, Peres K, Proust-Lima C. Individual dynamic prediction of clinical endpoint from large dimensional longitudinal biomarker history: a landmark approach. *BMC Med Res Methodol*. 2022;22(1):1. doi:10.1186/s12874-022-01660-3
27. Yang T, Yang Y, Jia Y, Li X. Dynamic prediction of hospital admission with medical claim data. *BMC Med Inform Decis Mak*. 2019;19(Suppl 1):18. doi:10.1186/s12911-019-0734-y
28. Lin J, Li K, Luo S. Functional survival forests for multivariate longitudinal outcomes: dynamic prediction of Alzheimer's disease progression. *Stat Methods Med Res*. 2021;30(1):99–111. doi:10.1177/0962280220941532
29. Fahimi F, Guo Y, Tong SC, et al. A vital signs telemonitoring programme improves the dynamic prediction of readmission risk in patients with heart failure. *AMIA Annu Symp Proc AMIA Symp*. 2020;2020:432–441.
30. Zeng Z, Tang X, Liu Y, He Z, Gong X. Interpretable recurrent neural network models for dynamic prediction of the extubation failure risk in patients with invasive mechanical ventilation in the intensive care unit. *BioData Min*. 2022;15(1):21. doi:10.1186/s13040-022-00309-7
31. Howell M, Wong G, Rose J, Tong A, Craig JC, Howard K. Patient preferences for outcomes after kidney transplantation: a best-worst scaling survey. *Transplantation*. 2017;101(11):2765–2773. doi:10.1097/TP.0000000000001793
32. Lenain R, Dantan E, Giral M, et al. External validation of the dynpg for kidney transplant recipients. *Transplantation*. 2021;105(2):396–403. doi:10.1097/TP.0000000000003209
33. Calster BV, van SM, van AW, Coemans M, Wynants L, Steyerberg EW. The enemies of reliable and useful clinical prediction models: a review of statistical and scientific challenges. *Annu. Rev. Stat. Appl*. 2026;13:465–492. doi:10.1146/annurev-statistics-042324-123749
34. Tong A, Gill J, Budde K, 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(8):1887–1896. doi:10.1097/TP.0000000000001774
35. Gordon EJ, Butt Z, Jensen SE, et al. Opportunities for shared decision making in kidney transplantation. *Am J Transplant off J Am Soc Transplant Am Soc Transpl Surg*. 2013;13(5):1149–1158. doi:10.1111/ajt.12195
36. Vandecasteele SJ, Kurella Tamura M. A patient-centered vision of care for ESRD: dialysis as a bridging treatment or as a final destination? *J Am Soc Nephrol JASN*. 2014;25(8):1647–1651. doi:10.1681/ASN.2013101082
37. Foucher Y, Meurette A, Daguin P, et al. A personalized follow-up of kidney transplant recipients using video conferencing based on a 1-year scoring system predictive of long term graft failure (TELEGRAFT study): protocol for a randomized controlled trial. *BMC Nephrol*. 2015;16:6. doi:10.1186/1471-2369-16-6
38. Jörres A, John S, Lewington A, et al. A European Renal Best Practice (ERBP) position statement on the Kidney Disease Improving Global Outcomes (KDIGO) Clinical Practice Guidelines on Acute Kidney Injury: part 2: renal replacement therapy. *Nephrol Dial Transplant off Publ Eur Dial Transpl Assoc - Eur Ren Assoc*. 2013;28(12):2940–2945. doi:10.1093/ndt/gft297
39. Evans JT, Evans JP, Walker RW, Blom AW, Whitehouse MR, Sayers A. How long does a Hip replacement last? A systematic review and meta-analysis of case series and national registry reports with more than 15 years of follow-up. *Lancet*. 2019;393(10172):647–654. doi:10.1016/S0140-6736(18)31665-9
40. Heazell AEP, Graham N, Parkes MJ, Wilkinson J. Dynamic prediction of pregnancy outcome after previous stillbirth or perinatal death: pilot study to establish proof-of-concept and explore method feasibility. *Ultrasound Obstet Gynecol off J Int Soc Ultrasound Obstet Gynecol*. 2024;64(5):613–619. doi:10.1002/uog.29104
41. Chen G, Chen Y, Shi Y, Mao Z, Lou J, Ma J. A dynamic prediction model for preeclampsia using the sFlt-1/PLGF ratio combined with multiple factors. *BMC Pregnancy Childbirth*. 2024;24(1):443. doi:10.1186/s12884-024-06627-4

42. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ*. 2024:385e078378. doi:10.1136/bmj-2023-078378.
43. Van Calster B, McLernon DJ, van Smeden M, et al. Calibration: the Achilles heel of predictive analytics. *BMC Med*. 2019;17(1):230. doi:10.1186/s12916-019-1466-7
44. Simon-Pimmel J, Foucher Y, Léger M, et al. Methodological quality of multivariate prognostic models for intracranial haemorrhages in intensive care units: a systematic review. *BMJ Open*. 2021;11(9):e047279. doi:10.1136/bmjopen-2020-047279
45. Damen JAAG, Hooft L, Schuit E, et al. Prediction models for cardiovascular disease risk in the general population: systematic review. *BMJ*. 2016;353:i2416. doi:10.1136/bmj.i2416
46. Navarro CLA, Damen JAA, Takada T, et al. Risk of bias in studies on prediction models developed using supervised machine learning techniques: systematic review. *BMJ*. 2021;375:n2281. doi:10.1136/bmj.n2281
47. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Ann Intern Med*. 2019;170(1):51–58. doi:10.7326/M18-1376
48. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Ann Intern Med*. 2019;170(1):W1–W33. doi:10.7326/M18-1377
49. Rezaeian M. Reducing publication bias in biomedical research: reviewing and registering protocol with a suitable journal. *J Clin Epidemiol*. 2016;69:248–249. doi:10.1016/j.jclinepi.2015.05.007
50. Andrinopoulou ER, Rizopoulos D, Takkenberg JJ, Lesaffre E. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Stat Methods Med Res*. 2017;26(4):1787–1801. doi:10.1177/0962280215588340
51. Suresh K, Taylor JMG, Spratt DE, Daignault S, Tsodikov A. Comparison of joint modeling and landmarking for dynamic prediction under an illness-death model. *Biom J Biom Z*. 2017;59(6):1277–1300. doi:10.1002/bimj.201600235
52. Yang M, Luo S, DeSantis S. Bayesian quantile regression joint models: inference and dynamic predictions. *Stat Methods Med Res*. 2019;28(8):2524–2537. doi:10.1177/0962280218784757
53. Desmée S, Mentré F, Veyrat-Follet C, Sébastien B, Guedj J. Nonlinear joint models for individual dynamic prediction of risk of death using Hamiltonian Monte Carlo: application to metastatic prostate cancer. *BMC Med Res Methodol*. 2017;17(1):105. doi:10.1186/s12874-017-0382-9
54. Li N, Liu Y, Li S, Elashoff RM, Li G. A flexible joint model for multiple longitudinal biomarkers and a time-to-event outcome: with applications to dynamic prediction using highly correlated biomarkers. *Biom. J*. 2021;63(8):1575–1586. doi:10.1002/uog.29104
55. Chabeau L, Rinder P, Desmée S, Giral M, Dantan E. A noise-tolerant inference procedure for Quasi-Monte Carlo likelihood estimation of a joint model for multiple longitudinal markers and competing risks. *Stat Med*. 2025;44(23–24):e70298. doi:10.1002/sim.70298

Clinical Epidemiology

Dovepress
Taylor & Francis Group

Publish your work in this journal

Clinical Epidemiology is an international, peer-reviewed, open access, online journal focusing on disease and drug epidemiology, identification of risk factors and screening procedures to develop optimal preventative initiatives and programs. Specific topics include: diagnosis, prognosis, treatment, screening, prevention, risk factor modification, systematic reviews, risk & safety of medical interventions, epidemiology & biostatistical methods, and evaluation of guidelines, translational medicine, health policies & economic evaluations. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use.

Submit your manuscript here: https://www.dovepress.com/clinical-epidemiology-journal