REVIEW   **Open Access**

Harnessing repeated measurements of predictor variables for clinical risk prediction: a review of existing methods

Lucy M. Bull[1,2]* , Mark Lunt[1], Glen P. Martin[3], Kimme Hyrich[1,4] and Jamie C. Sergeant[1,2]

Abstract

Background: Clinical prediction models (CPMs) predict the risk of health outcomes for individual patients. The majority of existing CPMs only harness cross-sectional patient information. Incorporating repeated measurements, such as those stored in electronic health records, into CPMs may provide an opportunity to enhance their performance. However, the number and complexity of methodological approaches available could make it difficult for researchers to explore this opportunity. Our objective was to review the literature and summarise existing approaches for harnessing repeated measurements of predictor variables in CPMs, primarily to make this field more accessible for applied researchers.

Methods: MEDLINE, Embase and Web of Science were searched for articles reporting the development of a multivariable CPM for individual-level prediction of future binary or time-to-event outcomes and modelling repeated measurements of at least one predictor. Information was extracted on the following: the methodology used, its specific aim, reported advantages and limitations, and software available to apply the method.

Results: The search revealed 217 relevant articles. Seven methodological frameworks were identified: time-dependent covariate modelling, generalised estimating equations, landmark analysis, two-stage modelling, joint-modelling, trajectory classification and machine learning. Each of these frameworks satisfies at least one of three aims: to better represent the predictor-outcome relationship over time, to infer a covariate value at a pre-specified time and to account for the effect of covariate change.

Conclusions: The applicability of identified methods depends on the motivation for including longitudinal information and the method's compatibility with the clinical context and available patient data, for both model development and risk estimation in practice.

Keywords: Longitudinal data, Clinical risk prediction, Dynamic prediction, Prediction models, Survival analysis, Repeated observations, Electronic health records, Personalised medicine, Time-dependent covariates, Joint models

* Correspondence: lucy.bull@manchester.ac.uk
[1]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
[2]Centre for Biostatistics, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK
Full list of author information is available at the end of the article

Bull *et al. Diagnostic and Prognostic Research* (2020) 4:9

Page 2 of 16

Background

Clinical prediction models (CPMs) aim to predict the risk of health outcomes such as disease onset, disease progression or likely outcomes of treatment [1]. Such predictions are based on available information about an individual at the time of prediction and can be used to inform patient care. This could be by offering preventative interventions to those predicted to be at high risk of an adverse outcome or relaxing the monitoring of those predicted to be at low risk. A clinical example of the former is the QRISK prediction tool currently used in primary care to estimate a patient's 10-year risk of cardiovascular disease (CVD) [2]. UK health guidelines advise that anyone with an estimated CVD risk of 10% or higher (from the QRISK model) should be prescribed a statin to reduce their risk of CVD onset [2, 3].

The majority of current CPMs use patient information from only a single time point to make predictions and fail to take advantage of longitudinal medical data, such as that available in electronic health records (EHRs). It has been hypothesized that repeated observations provide a predictive advantage over cross-sectional information as they capture change in individual patients over time and are less sensitive to measurement error [4, 5]. Furthermore, recent empirical reviews comparing longitudinal CPMs to the traditional cross-sectional ones provide some evidence that the overall predictive accuracy can be improved by incorporating the longitudinal patient information [5–9].

While an increasing number of CPMs are being developed using EHR data, a systematic review showed that less than 9% of identified CPMs exploited the time-varying nature of their predictor variables [6]. Therefore, although methods for longitudinal data analysis are well established, they appear to be under-utilised in the development of CPMs.

To the authors' knowledge, a broad review of available methods adopted for harnessing longitudinal data in binary or time-to-event CPMs has not yet been performed. Binary and time-to-event outcomes are of primary interest here as they are the most commonly reported amongst the prediction-modelling literature [6, 7]. Previous reviews have been restricted to simpler methods [8], methods most compatible to a particular clinical application [5, 9–11], or they have been restricted to the two commonly considered methods in the field of CPMs (i.e. joint models and landmark analysis, see Results for method description) [12–14]. The availability of a broad review could help the development of longitudinal CPMs and their potential use in practice.

Our primary objective was to review the literature and provide applied researchers with a comprehensive summary of existing approaches used for harnessing repeated measurements of predictors in CPMs. To address this objective, we sought to group identified methods based on their similarity and how they use repeated observations to enhance prediction, as well as outline their reported advantages and limitations. Our secondary objectives were to provide guidance on how to choose an appropriate method and to highlight opportunities for further methodological research.

Methods

Definitions and terminology

Within this review, longitudinal information is defined as repeated measurements through time of predictor variables. Predictor variables here are defined as measurable factors that are potentially predictive of health outcomes of interest. The terms `predictors' and `covariates' will also be used interchangeably for `predictor variables'. Note that the analytical methods discussed in this review are sometimes categorised under the term `dynamic prediction'. However, dynamic prediction can cover a broader range of aims and methods than those of interest here. In particular, the methods covered in this review are distinct to those for addressing calibration drift [15] or modelling disease state transitions [16].

Search strategy

The search strategy in Table 1 was designed to find peer-reviewed journal articles that described the development of a CPM for individual-level prediction of a binary or time-to-event outcome, and accounted for repeated measurements over time of at least one predictor variable during model development.

The search terms (Table 1) were entered into MEDLINE (via Ovid), Embase (via Ovid) and Web of Science. The search was restricted to peer-reviewed journal articles in English. Further details about any refinements specific to each database have been reported in Table 2. Duplicates were removed using automatic deduplication on both EndNote X8 and Mendeley Desktop.

Inclusion and exclusion criteria

A two-stage screening process was performed prior to full-text assessment for eligibility. Different sets of inclusion criteria were used to screen titles and abstracts, set A and set B respectively. Set B was also used for full-text assessment. Both sets of inclusion criteria are clearly stated, alongside the rule of inclusion, in Table 3. For an article to be taken through a stage of the screening process (or the full-text assessment), it must have satisfied the `rule of inclusion' (e.g. article titles that did not satisfy either criterion 1 or 2 alongside criterion 3 in set A were excluded from the review). If it was unclear whether an article satisfied the inclusion criteria in the relevant set, it was automatically brought forward to the next stage (i.e. to abstract screening or full-text assessment).

Table 1 Search strategy in Ovid format, as entered into MEDLINE and Embase

1 ((repeat* adj1 measure*) OR ``repeatedly-measured'' OR (repeat* adj1 observ*) OR ``repeatedly-observed'').ti,ab

2 ``time-series'' OR ``time-series'' OR (``longitudinal'' adj2 ``data'')).ti,ab

3 ``longitudinal'' adj3 (``survival'' OR ``binary'')) OR ((``longitudinal'' OR ``repeat'' OR ``discrete'') adj2 (``time-to-event'' OR (event* adj2 time*)))).ti,ab

4 ((time-depend* OR ``time-varying'' OR ``longitudinal'') adj1 (coefficient* OR variable* OR covariate* OR marker* OR factor* OR observ* OR measure* OR biomarker* OR model* OR predictor*)).ti,ab

5 ((predict* adj1 (accurac* OR ``power'' OR individual* OR ``future'' OR ``time-to-event'' OR (event* adj2 time*) OR ``binary'')) OR ``predictive ability'' OR ``predictive performance'').ti,ab

6 (((``predictive'' OR ``prediction'' OR ``prognostic'') adj1 (tool* OR scor* OR ``algorithm'' OR model* OR rule*))).ti,ab

7 ``predict'' or ``predicts'' or ``prediction'' or ``predicting'') adj2 (risk* OR ``outcome'' OR ``incidence'' OR ``time'' OR development* OR event* OR ``disease'' OR recurrence* OR ``progression'' OR ``severity'' OR ``achievement'' OR ``status'' OR ``application'')).ti,ab

8 ((``predict'' or ``predicts'' or ``prediction'' or ``predicting'') adj3 (``mortality'' OR ``survival'')).ti,ab

9 (``dynamic prediction'' OR ``dynamic predictions'' OR ``dynamic prognostic'' OR ``dynamic clinical prediction'').ti,ab

10 1 OR 2 OR 3 OR 4

11 5 OR 6 OR 7 OR 8

12 10 AND 11

13 12 OR 9

Table 2 Search limits specific to MEDLINE, Embase and Web of Science

Search limits			
Database	Ovid MEDLINE(R) and Epub Ahead of Print, In-process and Other Non-indexed Citations, Daily and Versions(R)	Ovid Embase	Web of Science
Search	Title and Abstract (.ti,ab)	Title and Abstract (.ti, ab)	Title, abstract, and keywords (TS=())
Dates	1946 to November 30, 2018	1974 to 2018 December 03	1900–2018
Publication status	No limit	Article-in-press, Embase status, In-Process status	No limit
Document type	Journal Article	Article, Article in press	Article
Language	English	English	English
Categories	No limit	No limit	Categories covering computer science, biology, healthcare, pharmacy, mathematics, statistics, biomedical engineering, psychology and probability.
Citation index	No limit	No limit	Social science, Science, Emerging sources.

Information extraction

The following information was extracted from relevant journal articles: the method for modelling longitudinal predictor variables, the aim of the method, the computer software used (if stated), the number and type of variables modelled longitudinally within the CPM, the clinical application and publication year. Publication years were extracted to provide a graphical overview of method usage over time. For any methods identified during the search, reported advantages, challenges and opportunities for future work regarding their application in CPMs were also extracted.

Results
Database search

The database search produced 10 615 results, which included 6960 unique peer-reviewed journal articles after

Table 3 Inclusion criteria used for the title, abstract and full-text screening

	Inclusion criteria set A (for titles)	Inclusion criteria set B (for abstracts and full-texts)
1	Development of a CPM.	Development of a multivariable CPM, which predicts a binary or time-to-event outcome.
2	Modelling techniques for longitudinal and survival/binary data.	CPM accounts for repeated measurements over time of at least one predictor variable.
3	Clinical application described or article published in a medical or biometric journal.	CPM has been developed for a binary or time-to-event out-come for an individual.
Inclusion rule	(1 OR 2) AND 3	1 AND 2 AND 3

the removal of duplicates, book chapters and conference proceedings. Following title screening, 752 articles remained in the review. The abstract screening and full-text assessment for eligibility left 247 and 217 articles respectively. The full screening process and reasons for exclusion have been described in Fig. 1. Additional file 1 lists all the articles included in the review.

Methodological review

The following terminology was identified within the review, which we here define to facilitate the understanding and comparison of methods described below:

landmark time, prediction time, horizon time, observation window and prediction window (see Fig. 2). Landmark time is equivalent to prediction time, which is the time that an individual's prognosis is being assessed. Horizon time is the end of the period that the prediction applies to. As an example, the QRISK models are developed to predict 10-year risk of cardiovascular disease, the horizon time is thus landmark time + 10 years [2]. Observation window refers to the period of time where a patient's covariate history can be observed for the purpose of inclusion into the CPM, which is always prior and up until the landmark/prediction time. Finally, the
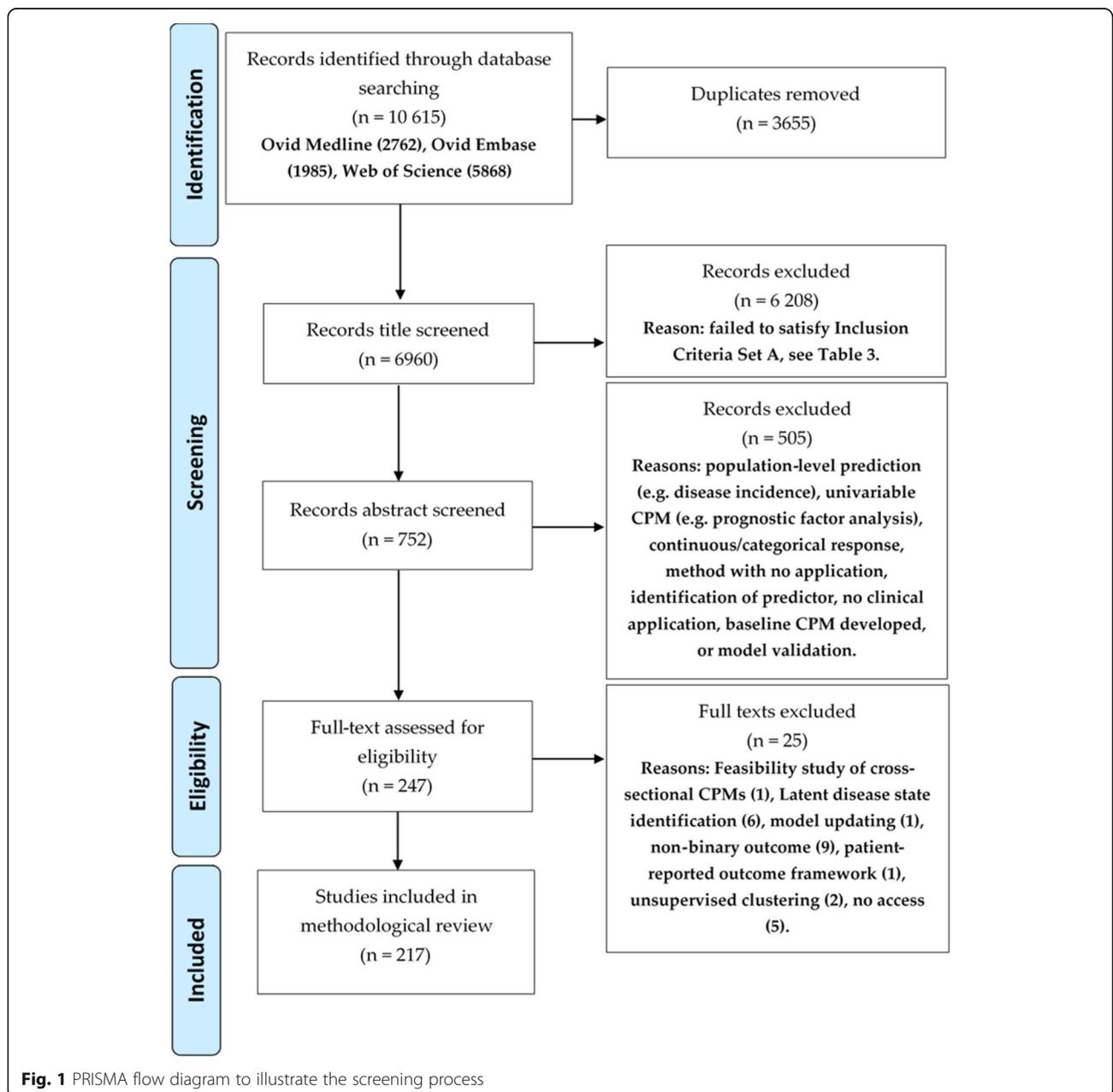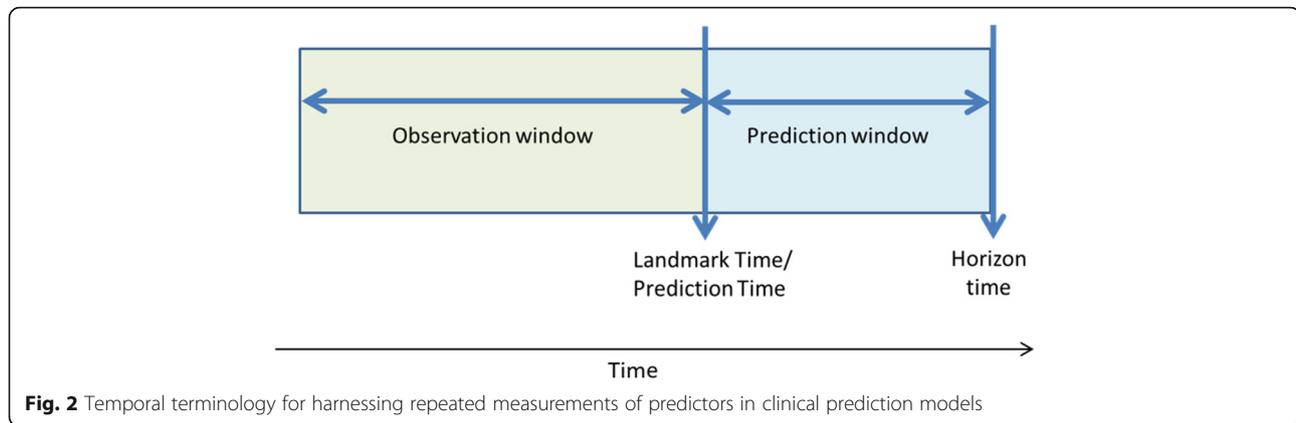

Fig. 1 PRISMA flow diagram to illustrate the screening process

Fig. 2 Temporal terminology for harnessing repeated measurements of predictors in clinical prediction models

prediction window is the time period between the land-mark time and horizon time.

From the included studies, three distinct methodological aims for harnessing repeatedly measured predictors in the development of CPMs were identified. All of the discovered methods satisfied one or more of these three methodological aims: (A1) to better represent the predictor-outcome relationship, (A2) to infer or predict a covariate value at a pre-specified time or (A3) to account for the effects of how a predictor changes over time. It is important to highlight that the content of this methodological review only covers methods reported in the identified literature via the database search, and that other valid approaches may exist but have not yet been applied in this field of clinical risk prediction.

Methods satisfying A1 tend to utilise repeated observations to represent a time-constant relationship, or better represent a time-varying relationship, between a predictor and the event of interest. Consequently, these methods often also allow for updated predictions through time. A2 is often pursued to either account for measurement error or random noise, or to impute missing data when measurements are irregularly-spaced. Methods for A3 are adopted when it is the behaviour of the covariate that is considered predictive of the event of interest.

In addition to identifying the three aims, the available methods were categorised into seven distinct frameworks: time-dependent covariate modelling (TDCM), generalised estimating equations (GEE), landmark analysis (LA), two-stage modelling (TSM), joint-modelling (JM), trajectory classification (TC) and machine learning (ML). All identified methods require subject-level longitudinal information on a study population for CPM development. The TSM, JM and TC frameworks (as well as some ML algorithms) can also harness a subject's repeated measurements at the time of prediction. Meanwhile, as stand-alone frameworks, the TDCM, GEE and LA frameworks only require a subject's most recent

observations (i.e. a maximum of one measurement for each predictor) at the time of prediction. Figure 3 provides an overview of framework adoption over the past decade, showing that the JM, TSM, LA and ML frameworks appear to be the most popular. GEE and TC frameworks are the least adopted frameworks.

Some of the above frameworks have been extended to harness `functional data', defined by Li and Luo [17] as data that 'provide information about curves, surfaces, or anything else varying over a continuum'. For our review, this includes functional data on both a one-dimensional time domain such as heart rate monitor data or electroencephalogram data, and on higher dimensional domains such as magnetic resonance imaging or positron emission tomography. Any extensions of methods for functional data will be specified below. Multivariate longitudinal methods for prediction tailored to functional data generated in a critical care setting were also outlined by Plate et al. in 2019 [11]. The remainder of this subsection provides a detailed description of each identified methodological framework in turn. The description includes how they use longitudinal information, their reported advantages and limitations, and their extensions. An overview of each framework, their corresponding available software and example clinical applications is reported in Table 4.

Time-dependent covariate modelling

The most prominent approach before 2009 was to include time-dependent covariates within a survival model [5, 18]. We refer to this technique as the TDCM framework, as it can be applied to various adaptations of Cox regression models. The TDCM framework allows for instantaneous risk estimates to be produced at any time within the observation window and prediction window conditional on survival up until that time, whilst harnessing an individual's most recent observations [5, 19–21]. Conceptually, this approach compares the most recent
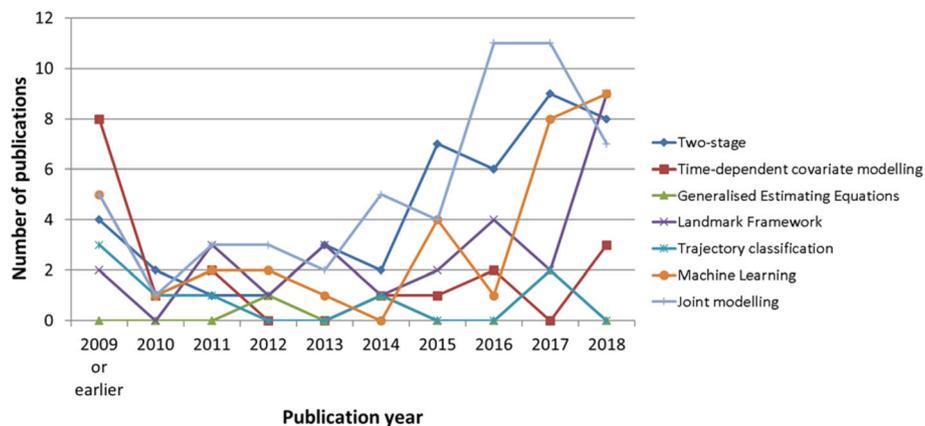
Fig. 3 Number of publications per year for each framework (*n* = 182), excluding methods in comparative reviews

covariate values for those still at risk just before each event time for those who have and have not experienced an event at that specific time. From there, the hazard function is updated over time and a time-constant effect between each covariate and the event of interest is estimated [5]. Therefore, TDCM falls under the first methodological aim (A1) and, as the timing of each event is required, can only handle time-to-event outcomes.

TDCM provides an advantage over baseline CPMs by enabling risk estimates to be updated during follow-up for new individuals, using their most recent covariate values [22]. Applying baseline CPMs to patient data collected during follow-up would lead to under-estimated risk predictions and over-estimated survival predictions [22]. However, TDCM has been heavily criticised throughout the literature for the following reasons. First, covariate values are assumed to be measured without error [5, 21]. Second, repeated covariate values over time are assumed to remain constant between data collection points [5, 21]. Third, correlations between and within subject measurements are not taken into account [21]. Finally, and most importantly, a time-dependent survival model is unable to predict into the future beyond the first change in the covariates [21, 23].

To elaborate on this final limitation, the challenge lies with the requirement of patient covariate values at the horizon time, as these are unknown for new individuals in practice. The simplest, and most common approach, to overcoming this final limitation is to use last observation carried forward (LOCF) from landmark time to horizon time [5]. This variation of TDCM has been employed, for example, to assess the prognosis of individuals with hepatocellular carcinoma at any stage of their disease using their most recent clinical information (Table 4) (103). The magnitude of the error introduced by the LOCF aspect of prediction for a new individual is usually dependent on

the prediction window size and the stability of predictor variables over time, with TDCM being argued as a valid approach for short-term prediction windows [24].

Extensions of TDCM can account for time-dependent effects of predictors [25, 26], and aim to minimise the error caused by the LOCF approach by including time since measurement as a predictor [27], or including aggregated summaries of covariates [26].

Generalised estimating equations

Similarly to the TDCM framework, the primary methodological aim of generalised estimating equations (GEE) is to utilise repeated observations from the same individual to better represent the association between the predictor variables and the event of interest. However, unlike the TDCM framework, GEE models account for within and between individual correlation, can directly harness repeated events per individual [28, 29], and can model either binary or survival outcomes. In general, GEE models are most suitable when the model development data violates the independence assumption and the model developer's primary interest is in the most accurate estimation of the predictor-outcome relationship.

More specifically for clinical risk prediction, GEEs have been employed to handle repetitions of cross-sectional patient information through time (both baseline and outcome information), which will here be referred to as `cycles' of patient information [28–30]. As an example, a patient may experience several cycles of the same treatment (e.g. chemotherapy) to treat their condition (e.g. cancer); therefore, multiple pre-treatment measurements and multiple post-treatment adverse outcomes per patient could be utilised to develop a CPM for predicting adverse outcomes from chemotherapy [28–30].

Table 4 Methodological frameworks available to enhance clinical prediction models using longitudinal information

Framework	Aim	Advantages	Limitations	Software	Extensions/variations	Examples
1. Time-dependent covariate modelling (TDCM)	A1	Allows for updated predictions over time, simple to apply in available software.	Assumes no measurement error, cannot predict the future, correlations[a] ignored, measurements assumed constant between time-points, requires complete predictors at event times.	Widely available (e.g. R, Stata, SAS).	Time-varying effects [25], time-since-measurement as a predictor [27], aggregated co-variate [26]. Quantile residual life regression [109].	Applied to assess the prognosis of patients with hepatocellular carcinoma, allowing for prediction at any stage of disease using their most recent information [110].
2. Generalised estimating equations (GEE)	A1	Allows for updated predictions over time, accounts for correlation[a], can adjust for patient clustering.	Ignores underlying trajectory, does not account for changes in at-risk population, and ignores time-dependency.	Widely available (e.g. R, stata, SAS). geepack package on R.		Employed to identify patients at high risk of adverse events after cancer therapy [28–30]. To account for repeated pre-therapy measurements and outcomes per individual though repeated treatment cycles.
3. Landmark analysis (LA)	A1	Avoids misspecification of underlying trajectory, only uses patient information prior to landmark time.	Ignores underlying covariate trajectory, often correlations[a] ignored, requires complete follow-up, and LOCF approach induces bias.	*dynpred* and *coxph* functions in R.	Competing risks [34, 41], recurrent events [36], combined with TSM [34, 38, 40], pseudo-observations [35, 41], cure fraction models [42].	Employed to predict relapse/death for those in leukaemia remission after transplant [34]. Landmark times 1, 6 and 12 months after bone marrow transplant [34]. Accounted for complications experienced by patients during follow-up.
4. Two-stage modelling (TSM)	A2 or A3	Simple to apply, flexible, can account for correlations[a], can handle irregularly spaced measurements.	Ignores model-specification error in the first-stage, first model cannot account for drop-out bias.	*refund* in MFPCA R package (FPC), *merlin* package on R (ME models).	Extends to TDCM [111], and LA [112], calibration error included in stage II [52, 60].	In conjunction with LA, TSM used to predict adverse events following endovascular abdominal aortic aneurysm repair [44]. ME models for aneurysm sac diameter change over time, with Cox model [44].
5. Joint-modelling (JM)	A1 and (A2 OR A3)	Address limitations of TSM framework, allows updated predictions over time, flexible.	Complex to implement, strong parametric assumptions, computationally intensive.	*JMbayes* or *JM* R packages. *lcmm* R package for JLCMs, *frailtypack* R package for JFMs.	Time-varying effects [13], Bayesian moving average [53, 107], various functions of random effects [13, 53, 73–75, 113], third JM to handle missing data and cure fraction models [66, 114–116].	Shared random effects JM employed for real-time predictions of prostate cancer recurrence [74]. A ME sub-model for log PSA over time, and a Cox sub-model used for the time-to-event outcome. Estimated using MCMC.
6. Trajectory classification (TC)	A1, A2 and A3	Accounts for correlation[a], irregularly-spaced measurements, informative processes, updated predictions, underlying trajectory.	Complex and computationally intensive for multivariate applications, parametric assumptions required for covariate trajectory.	*merlin* package on R (ME models), *Rstan* package (Gaussian processes).	Multivariate modelling using Gaussian processes [117]. Multivariate modelling and informative processes [76]	Employed to classify repeated measurements of hormone levels in early pregnancy to predict pregnancy success in the context of in vitro fertilization [54]. A nonlinear ME model for hormone levels over time. The binary outcome (pregnancy) modelled as an interaction.
7. Machine learning (ML)	A1 and A3	Few assumptions, handles high-dimensional data, can identify optimal trajectory characteristics.	Often predicts binary outcome, ignores right-censoring, large datasets required to avoid overfitting, often 'black box' algorithms.	*random-Forest* R package, *Adaboost* or *gbm* R packages for Boosting, *LibSVM* on R for SVMs.	Recurrent Neural Networks (RNNs) [98–101], Multiple measurements and time series SVM [118, 119], ME models and conditional inference trees [81].	RNNs employed to predict heart failure based on EHR data [98]. RNN identified patterns in previous and current diagnoses and quantified similarities with historic patients diagnosed with heart failure [98].

[a]Correlations between and within individuals
Abbreviations: *LOCF* last observation carried forward, *ME* mixed effect, *SVM* support vector machine, *MCMC* Markov chain Monte Carlo, *EHR* electronic health record, *JLCM* joint latent class model

Traditional logistic regression would not be able to utilise such patient information as it violates the independence assumption, that each observation (for an individual) is independent of other observations. An alternative model to handle cycles of patient information is a beta-geometric model that has been used to predict natural conception for women, after multiple cycles of a relevant procedure [31].

Landmark analysis

The landmark analysis (LA) framework derives separate cross-sectional CPMs for those still at risk at various landmark time points during follow-up. The flexibility in choice of model used to develop the CPM at each landmark time point allows for both binary and survival outcomes to be modelled under this framework. The methodological goal of this framework, which is similar to TDCM and based on conditional survival modelling [32], is to acknowledge that those who have survived for longer are more likely to have a better prognosis than those who have not [32]. Unlike TDCM, LA can use past or current information from new individuals to make predictions about their future [33]. The CPMs developed at each landmark time post-baseline can take into account the covariate history until the landmark time point, but a patient's most recent observation is employed in time-fixed CPMs [28]. For example, the LA framework has been adopted with binary time-dependent covariates to predict relapse or death for those in remission from leukaemia after a bone marrow transplant [34], as shown in Table 4. The CPMs developed post-baseline acknowledged whether the patient had or hadn't experienced complications since their surgery [34]. LA as a stand-alone framework utilises longitudinal information to account for change in an at-risk population when specifying the predictor-outcome relationship, and therefore satisfies the first methodological aim (A1).

For CPM development, it is common to merge all risk-sets (i.e. data required to develop a CPM at each landmark time) into a stacked dataset and fit just one model to the available data, including landmark time as an independent variable. This is often referred to as the `super landmark model' [34, 35]. Correlations between the within-subject observations can be accounted for using GEEs [35], and non-parametric time-varying coefficients can be modelled over landmark time points [36, 37].

The LA framework is a simple way to update risk predictions over time, without imposing too many assumptions on the available information, and it can handle a large number of time-dependent covariates [36–39]. Its simplicity may also lead it to being more robust to misuse in practice as it is straightforward to implement and interpret by the end-users [36, 39].

However, it appears that there is no general guidance on the choice of landmark times as they vary with each application. Examples include using quantiles of event times to capture the changes in the at-risk population [35] or using different follow-up appointment times in clinical practice [34]. Implementation can also be challenging for left-censored information, and routinely collected data with no defined baseline time-point [37]. CPMs developed in routinely collected data have used

age as the landmark time to overcome this barrier [38]. Furthermore, the LA framework carries the same limitations as any conditional survival model, which is the requirement of a large dataset, complete with long-term follow-up covariate and event information for each of the landmark time points [32].

Mixed-effects or auto-regressive time series models can also be used to capture a subject's covariate trajectory and predict the value of a covariate at each landmark time point [23, 34, 37, 38, 40]. Various survival models have also been applied to account for competing risks [41], recurrent events [36] and cure fraction models [42]. Thus, variations of the LA framework may fall under the second and third methodological aim (A2 and A3) if combined with the two-stage framework (discussed below).

Two-stage modelling

The two-stage modelling (TSM) framework considers the modelling processes for repeated measurements (longitudinal model) and outcome prediction (survival or binary-outcome model) separately. A parameter estimate from the longitudinal model is included as a fixed-time covariate in the survival (or binary-outcome) model. TSM as a stand-alone framework is often employed to satisfy the second or third methodological aim, depending on whether it is the behaviour of the predictor which is placed into the second model (A3), or the predicted value of a covariate at a pre-specified time (A2).

The key advantage of this approach is that it is computationally efficient, especially compared to joint modelling [21, 43]. However the two stages are performed separately and acknowledgment of any error in the estimation process for the longitudinal model is not carried forward into the outcome prediction model [5]. Therefore, any resulting predictions could appear too precise [5].

In the TSM framework, many different statistical models could be applied at each of the two stages in the CPM development process. The simplest and most common approaches applied for the first stage are to aggregate the repeated observations into a summary statistic or to fit a mixed-effect model (described below). Other examples include functional principal components and time series; these methods are described in Additional file 2. Examples of the survival or binary outcome models include Cox proportional hazards [44], logistic regression [45, 46] and partly conditional models [43].

Aggregated data In the aggregated data approach, all available covariate information up until prediction time is aggregated into a summary statistic. Examples include the use of the cumulative mean, rate of change, standard deviation or variance, coefficient of variation or the minimum/maximum value of available measurements

for each individual [2, 5, 8, 47–49]. The most common statistic for models developed on EHR data was the extreme (min/max) value of a predictor within a pre-specified observation window [50].

This approach attempts to minimise the effect of measurement error on individual risk predictions by summarising over the longitudinal trajectory. The clear advantages of this approach are the simplicity, lack of computational demand, reduced sensitivity to noisy data and ability to handle multiple repeatedly-measured predictors [9]. However, the unbiased estimates of the mean, standard deviation and variance assume no underlying trend, bias, or variability change in the process, which is unrealistic for most clinical data [5].

Mixed-effect models Mixed-effects (ME) models can also be referred to as random-effects, hierarchical or multi-level models. Their name derives from the idea that population-level information is used to support and enhance power for subject-level inference where individual measurements may be minimal. Population-level information is captured in fixed effects, and subject-level variations from the population are captured in random effects [9, 51]. These include linear mixed models [5, 51, 52] and generalised linear mixed models [9, 13, 40, 53].

ME models can be used to represent the longitudinal trajectory of a predictor variable over time, and may or may not include additional predictors for the longitudinal predictor outcome. The random effects from this model, which reflect individual-level rate of change or an inferred value of a predictor variable (at a pre-specified time), could be included into a cross-sectional CPM [46]. For example, a linear mixed model (LMM) has been employed to represent a patient's aneurysm sac diameter change over time [44]. Using this LMM, each new patient's aneurysm sac diameter and its rate of growth can be estimated at the landmark time (using their previous measurements). These values were then used as predictors in a Cox survival model to estimate their risk of an adverse event [44]. This clinical example also explores LA and TSM combined (Table 4) [44].

Although ME models are extremely flexible, challenges arise with correctly specifying a parametric trend over time and how to represent the individual rate of change in the final CPM. Linear models, quadratic growth curves [8, 52], fractional polynomials [44] and cubic splines [9, 13, 53] can be used to model the trend over time. Most CPM developers have adopted trends from previous literature in their specific field, but an appropriate model could also be found using data–driven techniques like the multiple fractional polynomial algorithm [44]. ME models can be extended to have t-distributed residuals with continuous outcomes to better handle outlier observations, and within-person correlations for

the repeated measurements [54, 55]. They can also be extended to account for sub-groups within a population using latent class methods [56].

Joint modelling

The joint-modelling (JM) framework addresses the limitations of the TSM framework by simultaneously estimating the longitudinal sub-model and the survival or binary outcome sub-model [13, 51, 52]. The term "joint model" more broadly refers to any number of statistical models estimated jointly, but here the literature focussed on jointly modelling a longitudinal model and a survival or binary outcome model. Similar to the TSM framework, a ME model was often employed for the covariate trajectory and a Cox proportional hazards model for a time-to-event outcome [57]. However, variations of the event prediction sub-model exist in the CPM literature such as binary event models [52, 58–61], parametric survival models [9], models for discrete-time data [9, 62, 63], models for competing risks [64], generalised linear models [58], and models for left-truncated data [65, 66]. Furthermore, the ME models could be for different types of data (e.g. functional data) [66, 67], modelling nonlinear functions [68], modelling nonparametric functions [69, 70] or linear quantile mixed models [71] depending on the clinical context.

Under all JM frameworks, the various sub-models typically involve shared random effects, or latent variables, whether they are continuous or discrete [5, 9]. For the purposes of clinical risk prediction, three different frameworks have been described: shared random effects (SRE) joint models, joint latent class models (JLCM) and joint frailty models for recurrent events (JFM). A clinical example of where a SRE joint model has been employed to predict prostate cancer recurrence has been highlighted in Table 4. Detailed descriptions of these sub-categories and their differences can be found in Additional file 2.

One challenge of using random effects in CPMs is estimating the risk of a future event for a new subject, as their random effects are unknown. To resolve this, random effects can be sampled from their posterior predictive distribution, which is based on the population-level distribution of random effects from the fitted joint model, the new subject's covariate values until the time of prediction, and conditional on the subject still being at risk at the time of prediction [72, 73].

A more popular choice is to employ the Monte Carlo simulation approach as it takes into account the uncertainty around the survival or event probability estimate [57, 72, 73]. Monte Carlo simulation is, conceptually, a procedure that repeatedly samples parameter estimates and random effects based on their estimated posterior distributions from the fitted joint model [72, 73]. A new

individual's random effects can be simulated from their posterior predictive distribution, as stated above [72, 73]. Repeatedly sampling from the posterior distributions allows for an empirical distribution around the estimated survival or event probability [73]. Monte Carlo simulation has been employed independently of the model estimation process and is reported to be computationally efficient in contrast to the joint model specification [74, 75].

Trajectory classification
Mixed-effect (ME) models have also been employed to classify longitudinal trajectories for binary events or categories; these methods have been grouped under the trajectory classification framework for this review [52, 76, 77]. The methods can alternatively be referred to as "longitudinal linear discriminant analysis" or "pattern mixture models" depending on their estimation and classification process.

In the TC framework for the prediction of binary events, the binary outcome value of 0 or 1 is seen as a latent class variable in the mixed-effect model. That is, the outcome variable interacts with all the predictors within the mixed-effects model (both fixed and random), which specifies the longitudinal trajectories. For CPM development, events are observed and so the parameters can be estimated, which is like modelling the event and non-event subjects separately. In practice, when the outcome is unknown, separate distributions of the longitudinal predictor values can be estimated based on the *event* and *non-event* ME model parameters, as well as the new subject's observed longitudinal values [78–80]. Both of these distributions can then be used to produce a discrimination score, which can later be used to classify the subject or to produce a posterior probability that the subject will experience the event [78–80].

A subject's risk or discriminant score can also be re-estimated when new information becomes available. Therefore, this framework satisfies both the first and third methodological aim (A1 and A3). To extend this approach to predict time-to-event outcomes, covariate trajectories may be classified into categories that can then be used as a predictor within a survival model. This extension can be performed under the two-stage modelling or joint-modelling framework, the latter approach is referred to as the joint latent class model in Additional file 2. The TC framework has also been extended to incorporate additional models to account for repeated binary events over time, and for informative processes [76].

Machine learning
The definition of the term *machine learning* can often be ambiguous as it covers a broad range of data-driven algorithms in the fields of statistics and computer science. For the purpose of prediction, ML algorithms extend from regression-based models (such as logistic regression) to more complex mathematical modelling (such as neural networks). Although methods identified under the field of machine learning are not independent of regression-based techniques described elsewhere, what distinguishes them is their algorithmic design.

In the longitudinal CPM literature, the terms machine learning, data mining and statistical learning have all been used to refer to the following algorithms: regularised logistic regression (RLR), elastic net (EN), random forests, gradient boosting, support vector machines (SVM), artificial neural networks, and naïve Bayes (NB). Additional algorithms have been categorised under `matching' algorithms for how they use repeated measurements for binary classification, where the conceptual interpretation is similar to that of the TC framework, please see Additional file 2 for further information. The majority of reported machine learning algorithms were employed to classify data for binary outcomes, with very limited attention on time-to-event outcomes [81–83].

Some of the algorithms stated above have been used in a TSM framework alongside other methods to capture the longitudinal covariate information, such as aggregation into summary statistics (RLR, EN, NB) [84–86], autoregressive time-series modelling (SVM) [87], Gaussian processes (SVM) [88] and temporal extraction [89]. The discussed methods are also often employed amongst an algorithm which performs variable selection, CPM development and performance assessment (internal validation) simultaneously [84].

The following subsections will provide a summary of the most common algorithms (temporal extraction, random forests, support vector machines, and artificial neural networks), and how they have been reported to incorporate longitudinal information in clinical risk prediction. All of the methods satisfy the third methodological aim (A3) and can account for the effect of covariate change on the event of interest.

Temporal extraction Temporal extraction can be used to define different change types in repeated observation data such as 'trends, statuses, and other complex time-related attributes' [89]. The temporal patterns over time can correspond to 13 different temporal operators: BEFORE, OVERLAPS, MEETS, STARTS, DURING, and their inverse relations, as well as EQUALS [89]. A simpler version of this technique only consists of increasing, decreasing or stationary temporal processes [89]. Variations of this conceptual idea exist in the machine learning literature such as time interval related patterns [90, 91], and sequential pattern mining [83]. These algorithms are usually embedded into an algorithmic framework which aims to match patterns over time in a

current patient to historical patient information and infer the probability of the outcome of interest [90, 92].

Random forests Random forests are composed of a set of low-correlated decision trees developed upon subsets of data generated via bootstrap sampling [84, 93, 94]. A single decision tree can incorporate nonlinear relationships and interactions of classifications with a simple representation of the data [85]. Both random forests and gradient boosting are reported to be advantageous when a CPM requires a large number of predictors [84]. Gradient boosting is an extension of the random forests that iteratively generates a sequence of decision trees based on the misclassification of a previous decision tree [95]. Although random forests have been used with longitudinal data, it is unclear whether any dependence on time or the ordering of measurements has been acknowledged as it has recently been described as a time-independent method [6].

Support vector machines Support vector machines (SVM) aim to maximise the distance between events and non-events in a high-dimensional space. SVMs explicitly divide the two domains with a linear or non-linear function, often estimated using a Kernel function [94, 96]. SVMs are less sensitive to outliers than standard logistic regression yet are more computationally intensive as they can harness high-dimensional covariate information [97].

While this approach has typically been combined (in a TSM framework) with aggregated data and time-series modelling techniques [87, 96, 97], SVMs have recently been employed as a one-stage approach for CPM development to harness repeated measurements in an observation window from EHR data [94]. It is thus implied that SVM can handle the longitudinal nature of predictor variables. However, time-dependency may have still been ignored as it has been reported that SVMs aggregate longitudinal information and ignore temporal relationships [98].

Artificial neural networks Artificial neural networks (ANNs) are a complex mathematical model designed to replicate the decision making process of the human brain. However, unlike the tree-based algorithms described above, the network is designed inside a black box, also known as the hidden layers [85]. ANNs are specifically able to process nonlinear relationships amongst dependent and independent variables whose relationships are complex, multidimensional and interactive [85].

Artificial neural networks may be able to respect the structure of longitudinal data, yet this is unclear amongst the literature. Descriptions of the hierarchical extension of ANNs explicitly state that time-dependent covariates can be incorporated into the network, although no explicit applications are suggested [82, 99]. Recurrent neural networks (RNN) are extensions of ANNs that have the ability to remember historical results, establish relationships across repeated measurements and acknowledge patterns over time [98, 100]. Unlike articles discussing other ML techniques, articles using RNNs have been explicit about the method's ability to harness high-dimensional data and tackle multivariate time-series problems for the prediction of a binary outcome [99, 101]. Clinically, RNNs were adopted to predict heart failure based on EHR data in 2018 [98], see Table 4.

Discussion

This methodological review has identified three ways in which available methods can utilise longitudinal information to enhance the performance of CPMs: (A1) to better represent the predictor-outcome relationship; (A2) to infer a covariate value at a pre-specified time and (A3) to account for the effects of predictor change over time. All identified methods have been categorised into seven methodological frameworks which use longitudinal information in different ways: time-dependent covariate modelling; generalised estimating equations; landmark analysis; two-stage modelling; joint-modelling; trajectory classification and machine learning. Four of these frameworks can harness subject-level repeated measurements at the time of prediction for a new individual, as well as subject-level longitudinal information on a study population for CPM development.

Recent reviews of available methods for modelling repeatedly measured predictor variables in the development of CPMs have focussed on evaluating their predictive advantage over cross-sectional CPMs [5, 8–12, 14, 21, 102–106]. The range of compared methods varies across reviews, although joint models are typically compared with other methods. The choice of reviewed methods has often been determined by the specific methodological problem, such as modelling a single longitudinal predictor [5, 8, 12, 21], modelling multiple irregularly measured predictor variables prior to a fixed landmark time [9, 10], modelling multivariate longitudinal data in a critical care setting [11] or handling large datasets with small numbers of events [9]. Other reviews have been designed for comparing a newly proposed method with other available methods [21] or comparing more complex approaches with simpler methods [105, 106]. Finally, some reviews focus solely on the comparison of the LA and JM frameworks, which are the most popular approaches [12, 14, 102, 104].

Problematically, previous reviews and studies often refer to methods using different names. For example, TDCM (with a Cox proportional hazards model) has

been reported as the 'last observation carried forward' approach for the way it handles repeated measurements at the time of prediction [5]. However, this is distinct from the 'most recent observation' approach discussed in another review, which refers to the application of a baseline CPM in EHR data [9]. Similarly, 'ordinary regression calibration' and 'risk-set regression calibration' methods have been defined as a sub-category of ME models as they have different assumptions for the random effects, yet these terms are not used elsewhere [5]. Therefore, we hope that this review will create a practical guide for researchers wishing to apply these methods, by providing a unified summary of the literature.

Welten et al. were the first reviewers to provide a set of available methods to address the methodological challenge and the practical implications of modelling repeated measurements for individual-level prediction [8]. However, the review focused on simple approaches (in the TSM framework), arguing that random-effects models are not appropriate for individual risk prediction in practice [8]. Nonetheless, it has now been argued that a new patient's random effects can be estimated through Monte Carlo simulation [72, 73]. Plate et al. also proposed a framework to facilitate the understanding and uptake of a variety of multivariate longitudinal methods for prediction in critical care in 2019 [11]. Despite the authors advocating the framework to be more widely applicable to EHR-style data, the proposed framework was specified prior to the systematic database search, whereas the structure of this review's output has been completely derived from the identified literature.

As the scope of the current methodological review was not restricted to a particular clinical application or a particular set of methods, a key strength of this study is its ability to provide a broader overview of available methodology, directly compare how methods use longitudinal information, and highlight some key considerations for applied researchers when choosing an appropriate method. These key considerations include, but are not restricted to, the type and amount of information available at the time of prediction (including the number and type of longitudinal predictors), how the CPM can benefit from the longitudinal information and what is known a priori to model development (i.e. imposed model assumptions).

During this review, considerations for future methodological research were also identified. The following aspects of CPM development were often overlooked within the current literature: sample size requirements, the handling of missing or irregularly-spaced data, effectively summarising longitudinal information, model validation (and avoiding statistical overfitting) and finally, how to quantify the improvement in predictive accuracy when incorporating repeated measurements. To elaborate, irregularly spaced measurements cannot be directly modelled using some methods, and so require additional imputation methods. Potential algorithms are emerging to choose the best way to summarise longitudinal trajectories in joint models [107], but there is limited discussion elsewhere around variable selection in a longitudinal context. Model validation techniques remain similar to those for cross-sectional CPMs where applicable, and the quantification of predictive improvement is often performed using differences in C-index which lacks clinical interpretability [108]. We recommend future research in each of these areas.

The limitations of our study should be kept in mind when interpreting its results. First, the systematic search employed to identify available methodology was designed using free-text. The evolving nature of this research space has resulted in a lack of uniformity in language when referring to repeated measurements of predictor variables and dynamic prediction amongst the literature. Second, the screening was performed by one author (LB) which may have introduced subjectivity and bias into the screening process. Third, while aggregate details of available software are provided in Table 4, the frequency and the level of detail of software reporting in the identified literature were not assessed.

The first limitation may have resulted in the systematic search missing some methods. To minimise this risk, the search in Table 1 was designed to cover all methodological purposes of longitudinal data for clinical risk prediction, including prediction of population-level change, the identification of predictors and methods to address calibration drift. Furthermore, an initial title screening with a much broader search strategy was performed to identify further relevant articles from which key words and terminology could be extracted. To reduce the risk of subjectivity as a result of the second limitation, the reviewer remained cautious about articles where it was unclear whether they fit the inclusion criteria and carried them forward to the next round of screening. Despite a general lack of detail in the literature in the reporting of available software for discussed methods, all reported available software has been included in Table 4. Such information will be useful for the implementation of identified methods and can provide an indication of where software may not have been well-reported.

Having compared how methods use longitudinal information, summarised their reported advantages and disadvantages and grouped them based on methodological approach, we hope to facilitate the understanding of a broad and complex research domain. The findings from this review consolidate the message from previous

Bull *et al. Diagnostic and Prognostic Research*   (2020) 4:9

Page 13 of 16

reviews, that there is no straight-forward approach to developing a longitudinal CPM. However, to reiterate, the choice of methods is substantially reduced by the following considerations: the type and amount of information available at the time of prediction (including the number and type of longitudinal predictors), how the CPM can benefit from the longitudinal information and the validity of any assumptions for a specific application.

Conclusions

We have grouped methods available for incorporating repeatedly measured predictor variables into the development of a CPM, identified their methodological aims and discussed their reported advantages and limitations. In addition, amongst the literature we found some key considerations for CPM development and identified opportunities for further methodological research. Most importantly, however, our review has identified seven methodological frameworks which offer a wide range of ways in which longitudinal information can enhance CPMs by improving the representation of a predictor-outcome relationship, updating predictions during follow-up, inferring covariate values, or accounting for the effect of how a predictor variable changes over time.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s41512-020-00078-z.

Additional file 1: Table 1 Reference List. Title of data: Journal articles included in the review, ordered alphabetically by first author surname. Description of data: A list of all articles included in the review, including author names, title of publication, journal of publication, volume, pages and DOI number.

Additional file 2: Additional Results. Title of data: Additional methods identified in the review. Description of data: Further details about methodology discussed within the review.

Abbreviations

CPM: Clinical prediction model; CVD: Cardiovascular disease; EHR: Electronic health record; TDCM: Time-dependent covariate modelling; GEE: Generalised estimating equations; LA: Landmark analysis; TSM: Two-stage modelling; JM: Joint modelling; TC: Trajectory classification; ML: Machine learning; LOCF: Last observation carried forward; ME: Mixed effect; LMM: Linear mixed model; SRE: Shared random effect; JLCM: Joint latent class model; JFM: Joint frailty model; RLR: Regularised logistic regression; EN: Elastic net; SVM: Support vector machine; ANN: Artificial neural network; RNN: Recurrent neural network; MCMC: Markov chain Monte Carlo

Acknowledgements

Not applicable.

Authors' contributions

LMB contributed to the design of the study, conducted the literature search, reviewed and synthesised the literature and wrote the manuscript. ML contributed to the design of the study and the interpretation of findings, and reviewed/edited the manuscript. GPM contributed to the design of the study and the interpretation of findings, and reviewed/edited the manuscript. KH contributed to the interpretation of findings and reviewed/edited the manuscript. JS contributed to the design of the study and the interpretation of findings, and reviewed/edited the manuscript. All authors read and approved the final manuscript.

Availability of data and materials

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Centre for Epidemiology Versus Arthritis, Centre for Musculoskeletal Research, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK. [2]Centre for Biostatistics, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK. [3]Division of Informatics, Imaging and Data Science, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK. [4]National Institute for Health Research Manchester Biomedical Research Centre, Manchester University NHS Foundation Trust, Manchester Academic Health Science Centre, Manchester, UK.

References

1. Steyerberg EW. Introduction. In: Steyerberg EW, editor. Clinical prediction models: a practical approach to development, validation, and updating. Cham: Springer International Publishing; 2019. p. 1–11.
2. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ (Online). 2017;357:j2099.
3. National Institute for Health and Care Excellence. Cardiovascular disease: risk assessment and reduction, including lipid modification [Online]. London: NICE; 2014. [updated Sept 2016, Cited April 2020] (Clinical guideline [CG181]). Available from: https://www.nice.org.uk/guidance/cg181/.
4. Steyerberg EW, Moons KGM, van der Windt DA, Hayden JA, Perel P, Schroter S, et al. Prognosis research strategy (PROGRESS) 3: prognostic model research. PLoS Med. 2013;10(2):e1001381.
5. Sweeting MJ, Barrett JK, Thompson SG, Wood AM. The use of repeated blood pressure measures for cardiovascular risk prediction: a comparison of statistical models in the ARIC study. Stat Med. 2017;36:4514–28.
6. Goldstein BA, Navar AM, Pencina MJ, Ioannidis JPA. Opportunities and challenges in developing risk prediction models with electronic health records data: a systematic review. J Am Med Inform Assoc. 2016;24(1):198–208.
7. Bouwmeester W, Zuithoff NPA, Mallett S, Geerlings MI, Vergouwe Y, Steyerberg EW, et al. Reporting and methods in clinical prediction research: a systematic review. PLoS Med. 2012;9(5):1–12.
8. Welten M, de Kroon MLA, Renders CM, Steyerberg EW, Raat H, Twisk JWR, et al. Repeatedly measured predictors: a comparison of methods for prediction modeling. Diagn and Progn Res. 2018;2(1):5.
9. 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;36:2750–63.
10. Rahimian F, Salimi-Khorshidi G, Payberah AH, Tran J, Ayala Solares R, Raimondi F, et al. Predicting the risk of emergency admission with machine

learning: development and validation using linked electronic health records. PLoS Med. 2018;15:e1002695.

11. Plate JDJ, van de Leur RR, Leenen LPH, Hietbrink F, Peelen LM, Eijkemans MJC. Incorporating repeated measurements into prediction models in the critical care setting: a framework, systematic review and meta-analysis. BMC Med Res Methodol. 2019;19(1):199.

12. Rizopoulos D, Molenberghs G, Lesaffre EMEH. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. Biom J. 2017;59:1261–76.

13. Andrinopoulou ER, Eilers PHC, Takkenberg JJM, Rizopoulos D. Improved dynamic predictions from joint models of longitudinal and survival data with time-varying effects using P-splines. Biometrics. 2018;74:685–93.

14. 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. 2017;59:1277–300.

15. Jenkins DA, Sperrin M, Martin GP, Peek N. Dynamic models to predict health outcomes: current status and methodological challenges. Diagn and Progn Res. 2018;2(1):23.

16. Goyal D, Tjandra D, Migrino RQ, Giordani B, Syed Z, Wiens J. Characterizing heterogeneity in the progression of Alzheimer's disease using longitudinal clinical and neuroimaging biomarkers. Alzheimers Dement (Amst). 2018;10: 629–37.

17. 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. 2017:962280217722177.

18. Wong J, Taljaard M, Forster AJ, Escobar GJ, Van Walraven C. Addition of time-dependent covariates to a survival model significantly improved predictions for daily risk of hospital death. J Eval Clin Pract. 2013;19:351–7.

19. Su J, Barbera L, Sutradhar R. Do repeated assessments of performance status improve predictions for risk of death among patients with cancer? A population-based cohort study. Palliat Med. 2015;29:547–53.

20. Wang Y, Chen W, Heard K, Kollef MH, Bailey TC, Cui Z, et al. Mortality prediction in ICUs using a novel time-slicing Cox regression method. AMIA Annu Symp Proc. 2015;2015:1289–95.

21. Maziarz M, Heagerty P, Cai TX, Zheng YY. On longitudinal prediction with time-to-event outcome: comparison of modeling options. Biometrics. 2017; 73:83–93.

22. Boberg KM, Rocca G, Egeland T, Bergquist A, Broome U, Caballeria L, et al. Time-dependent Cox regression model is superior in prediction of prognosis in primary sclerosing cholangitis. Hepatology. 2002;35:652–7.

23. van Houwelingen HC, Putter H. Dynamic predicting by landmarking as an alternative for multi-state modeling: an application to acute lymphoid leukemia data. Lifetime Data Anal. 2008;14:447–63.

24. Christensen E, Altman DG, Neuberger J, De Stavola BL, Tygstrup N, Williams R, et al. Updating prognosis in primary biliary cirrhosis using a time-dependent Cox regression model. PBC1 and PBC2 trial groups. Gastroenterology. 1993;105:1865–76.

25. Hofner B, Kneib T, Hartl W, Kuchenhoff H. Building Cox-type structured hazard regression models with time-varying effects. Stat Model. 2011;11:3–24.

26. Reinikainen J, Laatikainen T, Karvanen J, Tolonen H. Lifetime cumulative risk factors predict cardiovascular disease mortality in a 50-year follow-up study in Finland. Int J Epidemiol. 2015;44:108–16.

27. de Bruijne MHJ, le Cessie S, Kluin-Nelemans HC, van Houwelingen HC. On the use of Cox regression in the presence of an irregularly observed time-dependent covariate. Stat Med. 2001;20:3817–29.

28. Dranitsaris G, Vincent MD, Yu J, Huang L, Fang F, Lacouture ME. Development and validation of a prediction index for hand-foot skin reaction in cancer patients receiving sorafenib. Ann Oncol. 2012;23:2103–8.

29. Dranitsaris G, Molassiotis A, Clemons M, Roeland E, Schwartzberg L, Dielenseger P, et al. The development of a prediction tool to identify cancer patients at high risk for chemotherapyinduced nausea and vomiting. Ann Oncol. 2017;28:1260–7.

30. Dranitsaris G, Lacouture ME. Development of prediction tools for diarrhea and rash in breast cancer patients receiving lapatinib in combination with capecitabine. Breast Cancer Res Treat. 2014;147:631–8.

31. Van Eekelen R, Scholten I, Tjon-Kon-Fat RI, Van Der Steeg JW, Steures P, Hompes P, et al. Natural conception: repeated predictions overtime. Hum Reprod. 2017;32:346–53.

32. Hieke S, Kleber M, Konig C, Engelhardt M, Schumacher M. Conditional survival: a useful concept to provide information on how prognosis evolves over time. Clin Cancer Res. 2015;21:1530–6.

33. Putter H, van Houwelingen HC. Understanding landmarking and its relation with time-dependent Cox regression. Stat Biosci. 2017;9(2):489–503.

34. Cortese G, Gerds TA, Andersen PK. Comparing predictions among competing risks models with time-dependent covariates. Stat Med. 2013;32: 3089–101.

35. Grand MK, de Witte TJM, Putter H. Dynamic prediction of cumulative incidence functions by direct binomial regression. Biom J. 2018;60:734–47.

36. Musoro JZ, Struijk GH, Geskus RB, Ten Berge I, Zwinderman AH. Dynamic prediction of recurrent events data by landmarking with application to a follow-up study of patients after kidney transplant. Stat Methods Med Res. 2018;27:832–45.

37. van Houwelingen HC. Dynamic prediction by landmarking in event history analysis. Scand J Stat. 2007;34:70–85.

38. Keogh RH, Seaman SR, Barrett JK, Taylor-Robinson D, Szczesniak R. Dynamic prediction of survival in cystic fibrosis: a landmarking analysis using UK patient registry data. Epidemiology. 2018.

39. Parast L, Cai T. Landmark risk prediction of residual life for breast cancer survival. Stat Med. 2013;32:3459–71.

40. Paige E, Barrett J, Stevens D, Keogh RH, Sweeting MJ, Nazareth I, et al. Landmark models for optimizing the use of repeated measurements of risk factors in electronic health records to predict future disease risk. Am J Epidemiol. 2018;187:1530–8.

41. Nicolaie MA, van Houwelingen JC, de Witte TM, Putter H. Dynamic pseudo-observations: a robust approach to dynamic prediction in competing risks. Biometrics. 2013;69:1043–52.

42. Shi HL, Yin GS. Landmark cure rate models with time-dependent covariates. Stat Methods Med Res. 2017;26:2042–54.

43. Konerman MA, Brown M, Zheng Y, Lok AS. Dynamic prediction of risk of liver-related outcomes in chronic hepatitis C using routinely collected data. J Viral Hepat. 2016;23:455–63.

44. Grootes I, Barrett JK, Ulug P, Rohlffs F, Laukontaus SJ, Tulamo R, et al. Predicting risk of rupture and rupture-preventing reinterventions following endovascular abdominal aortic aneurysm repair. Br J Surg. 2018;105:1294–304.

45. Maruyama N, Takahashi F, Takeuchi M. Prediction of an outcome using trajectories estimated from a linear mixed model. J Biopharm Stat. 2009;19:779–90.

46. Zeng C, Ellis JL, Steiner JF, Shoup JA, McQuillan DB, Bayliss EA. Assessment of morbidity over time in predicting health outcomes. Med Care. 2014; 52(Suppl 3):S52–9.

47. Wang CY, Baldwin LM, Saver BG, Dobie SA, Green PK, Cai Y, et al. The contribution of longitudinal comorbidity measurements to survival analysis. Med Care. 2009;47:813–21.

48. McCudden C, Akbari A, White CA, Biyani M, Hiremath S, Brown PA, et al. Individual patient variability with the application of the kidney failure risk equation in advanced chronic kidney disease. PLoS One. 2018;13.

49. Cho IJ, Sung JM, Chang HJ, Chung N, Kim HC. Incremental value of repeated risk factor measurements for cardiovascular disease prediction in middle-aged korean adults: results from the NHIS-HEALS (National Health Insurance System-National Health Screening Cohort). Circ-Cardiovasc Qual Outcomes. 2017;10. https://doi.org/10.1161/CIRCOUTCOMES.117.004197.

50. Mubeen AM, Asaei A, Bachman AH, Sidtis JJ, Ardekani BA, Alzheimer's Dis Neuroimaging I. A six-month longitudinal evaluation significantly improves accuracy of predicting incipient Alzheimer's disease in mild cognitive impairment. J Neuroradiol. 2017;44:381–7.

51. Qiu F, Stein CM, Elston RC, TBRU TBRU. Joint modeling of longitudinal data and discrete-time survival outcome. Stat Methods Med Res. 2016;25:1512–26.

52. Albert PS. A linear mixed model for predicting a binary event from longitudinal data under random effects misspecification. Stat Med. 2012;31:143–54.

53. Andrinopoulou ER, Rizopoulos D, Takkenberg JJM, Lesaffre E. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. Stat Methods Med Res. 2017;26:1787–801.

54. DelaCruz R, Fuentes C, Meza C, Lee DJ, Arribas-Gil A, De la Cruz R, et al. Predicting pregnancy outcomes using longitudinal information: a penalized splines mixed-effects model approach. Stat Med. 2017;36:2120–34.

55. Lin TI, Lee JC. A robust approach to t linear mixed models applied to multiple sclerosis data. Stat Med. 2006;25:1397–412.

56. Fojo AT, Musliner KL, Zandi PP, Zeger SL. A precision medicine approach for psychiatric disease based on repeated symptom scores. J Psychiatr Res. 2017;95:147–55.

57. Proust-Lima C, Sene M, Taylor JMG, Jacqmin-Gadda H. Joint latent class models for longitudinal and time-to-event data: a review. Stat Methods Med Res. 2014;23:74–90.

58. Horrocks J, van Den Heuvel MJ. Prediction of pregnancy: a joint model for longitudinal and binary data. Bayesian Anal. 2009;4:523–38.

59. Hwang Y-T, Wang C-C, Wang CH, Tseng Y-K, Chang Y-J. Joint model of multiple longitudinal measures and a binary outcome: An application to predict orthostatic hypertension for subacute stroke patients. Biom J. 2015; 57(4):661–75.

60. Kim S, Albert PS. A class of joint models for multivariate longitudinal measurements and a binary event. Biometrics. 2016;72:917–25.

61. Proust-Lima C, Letenneur L, Jacqmin-Gadda H. A nonlinear latent clss model for joint analysis of multivariate longitudinal data and a binary outcome. Stat Med. 2007;26:2229–45.

62. Rondeau V, Mauguen A, Laurent A, Berr C, Helmer C. Dynamic prediction models for clustered and interval-censored outcomes: investigating the intra-couple correlation in the risk of dementia. Stat Methods Med Res. 2017;26:2168–83.

63. McLain AC, Sundaram R, Buck Louis GM. Joint analysis of longitudinal and survival data measured on nested timescales by using shared parameter models: an application to fecundity data. J R Stat Soc: Ser C: Appl Stat. 2015; 64:339–57.

64. Musoro JZ, Zwinderman AH, Abu-Hanna A, Bosman R, Geskus RB. Dynamic prediction of mortality among patients in intensive care using the sequential organ failure assessment (SOFA) score: a joint competing risk survival and longitudinal modeling approach. Statistica Neerlandica. 2018;72: 34–47.

65. Piccorelli AV, Schluchter MD. Jointly modeling the relationship between longitudinal and survival data subject to left truncation with applications to cystic fibrosis. Stat Med. 2012;31:3931–45.

66. Piulachs X, Alemany R, Guillen M, Rizopoulos D. Joint models for longitudinal counts and left-truncated time-to-event data with applications to health insurance. Sort-Statistics and Operations Research Transactions. 2017;41:347–71.

67. Li K, Luo S. Bayesian functional joint models for multivariate longitudinal and time-to-event data. Comput Stat Data Anal. 2019;129:14–29.

68. Desmée S, Mentré F, Veyrat-Follet C, Sébastien B, Guedj J, Desmee S, et al. 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;17:105.

69. Szczesniak R, Su W, Clancy JP. Dynamics of disease progression and gastrostomy tube placement in children and adolescents with cystic fibrosis: application of joint models for longitudinal and time-to-event data. Internal medicine review (Washington, DC : Online). 2016;2.

70. Li K, Furr-Stimming E, Paulsen JS, Luo S. Dynamic prediction of motor diagnosis in Huntington's disease using a joint modeling approach. J Huntingtons Dis. 2017;6:127–37.

71. Yang M, Luo S, DeSantis S. Bayesian quantile regression joint models: inference and dynamic predictions. Stat Methods Med Res. 2018.

72. Krol A, Tournigand C, Michiels S, Rondeau V. Multivariate joint frailty model for the analysis of nonlinear tumor kinetics and dynamic predictions of death. Stat Med. 2018;37:2148–61.

73. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. Biometrics. 2011;67:819–29.

74. Taylor JM, Park Y, Ankerst DP, Proust-Lima C, Williams S, Kestin L, et al. Real-time individual predictions of prostate cancer recurrence using joint models. Biometrics. 2013;69:206–13.

75. Yang L, Yu M, Gao S. Prediction of coronary artery disease risk based on multiple longitudinal biomarkers. Stat Med. 2016;35:1299–314.

76. Coley RY, Fisher AJ, Mamawala M, Carter HB, Pienta KJ, Zeger SL. A Bayesian hierarchical model for prediction of latent health states from multiple data sources with application to active surveillance of prostate cancer. Biometrics. 2017;73:625–34.

77. Huang X, Yan F, Ning J, Feng Z, Choi S, Cortes J. A two-stage approachee for dynamic prediction of time-to-event distributions. Stat Med. 2016;35:2167–82.

78. Liu DP, Albert PS. Combination of longitudinal biomarkers in predicting binary events. Biostatistics. 2014;15:706–18.

79. Lukasiewicz E, Gorfine M, Neumann AU, Freedman LS. Combining longitudinal discriminant analysis and partial area under the ROC curve to predict non-response to treatment for hepatitis C virus. Stat Methods Med Res. 2011;20:275–89.

80. Brant LJ, Sheng SL, Morrell CH, Zonderman AB. Data from a longitudinal study provided measurements of cognition to screen for Alzheimer's disease. J Clin Epidemiol. 2005;58:701–7.

81. Wallace ML, Anderson SJ, Mazumdar S, Kong L, Mulsant BH. Incorporating temporal features of repeatedly measured covariates into tree-structured survival models. Biom J. 2012;54:181–96.

82. Ohno-Machado L, Walker MG, Musen MA. Hierarchical neural networks for survival analysis. Medinfo. 1995;8(Pt 1):828–32.

83. Smedley NF, Ellingson BM, Cloughesy TF, Hsu W. Longitudinal patterns in clinical and imaging measurements predict residual survival in glioblastoma patients. Sci Rep. 2018;8.

84. Yahya N, Ebert MA, Bulsara M, House MJ, Kennedy A, Joseph DJ, et al. Statistical-learning strategies generate only modestly performing predictive models for urinary symptoms following external beam radiotherapy of the prostate: a comparison of conventional and machine-learning methods. Med Phys. 2016;43:2040–52.

85. Lu HY, Li TC, Tu YK, Tsai JC, Lai HS, Kuo LT. Predicting long-term outcome after traumatic brain injury using repeated measurements of Glasgow coma scale and data mining methods. J Med Syst. 2015;39.

86. Ng K, Steinhubl SR, deFilippi C, Dey S, Stewart WF. Early detection of heart failure using electronic health records practical implications for time before diagnosis, data diversity, data quantity, and data density. circulation-cardiovascular quality and outcomes. 2016;9:649-58.

87. Minhas S, Khanum A, Riaz F, Khan SA, Alvi A. Alzheimers dis neuroimaging l. predicting progression from mild cognitive impairment to Alzheimer's disease using autoregressive modelling of longitudinal and multimodal biomarkers. IEEE J Biomed Health Inform. 2018;22:818–25.

88. Soguero-Ruiz C, Fei WM, Jenssen R, Augestad KM, Alvarez JL, Jimenez IM, et al. Data-driven temporal prediction of surgical site infection. AMIA Annu Symp Proc. 2015;2015:1164–73.

89. Cheng LC, Hu YH, Chiou SH. Applying the temporal abstraction technique to the prediction of chronic kidney disease progression. J Med Syst. 2017; 41(5):85.

90. Moskovitch R, Choi H, Hripcsak G, Tatonetti N. Prognosis of clinical outcomes with temporal patterns and experiences with one class feature selection. IEEE/ACM Trans Comput Biol Bioinform. 2017;14:555–63.

91. Moskovitch R, Polubriaginof F, Weiss A, Ryan P, Tatonetti N. Procedure prediction from symbolic electronic health records via time intervals analytics. J Biomed Inform. 2017;75:70–82.

92. Ping XO, Tseng YJ, Lin YP, Chiu HJ, Lai FP, Liang JD, et al. A multiple measurements case-based reasoning method for predicting recurrent status of liver cancer patients. Comput Ind. 2015;69:12–21.

93. Walsh CG, Ribeiro JD, Franklin JC. Predicting suicide attempts in adolescents with longitudinal clinical data and machine learning. J Child Psychol Psychiatry. 2018;59:1261–70.

94. An S, Malhotra K, Dilley C, Han-Burgess E, Valdez JN, Robertson J, et al. Predicting drug-resistant epilepsy - a machine learning approach based on administrative claims data. Epilepsy Behav. 2018;89:118–25.

95. Koyner JL, Carey KA, Edelson DP, Churpek MM. The development of a machine learning inpatient acute kidney injury prediction model. Crit Care Med. 2018;46:1070–7.

96. Lacson R. Predicting hemodialysis mortality utilizing blood pressure trends. AMIA 2008;Annual Sym:369-73.

97. Ajemba PO, Ramirez L, Durdle NG, Hill DL, Raso VJ. A support vectors classifier approach to predicting the risk of progression of adolescent idiopathic scoliosis. IEEE Trans Inf Technol Biomed. 2005;9:276–82.

98. Jin B, Che C, Liu Z, Zhang SL, Yin XM, Wei XP. Predicting the risk of heart failure with EHR sequential data modeling. Ieee Access. 2018;6:9256–61.

99. Parmanto B, Doyle HR. Recurrent neural networks for predicting outcomes after liver transplantation: representing temporal sequence of clinical observations. Methods Inf Med. 2001;40:386–91.

100. Ceccarelli F, Sciandrone M, Perricone C, Galvan G, Morelli F, Vicente LN, et al. Prediction of chronic damage in systemic lupus erythematosus by using machine-learning models. PLoS One. 2017;12.

101. Orchard P, Agakova A, Pinnock H, Burton CD, Sarran C, Agakov F, et al. Improving prediction of risk of hospital admission in chronic obstructive pulmonary disease: application of machine learning to telemonitoring data. J Med Internet Res. 2018;20.

102. 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. 2018. https://doi.org/10.1177/0962280218811837.

103. Kennedy CE, Turley JP. Time series analysis as input for clinical predictive modeling: modeling cardiac arrest in a pediatric ICU. Theor Biol Med Model. 2011;8.

104. Li L, Luo S, Hu B, Greene T. Dynamic prediction of renal failure using longitudinal biomarkers in a cohort study of chronic kidney disease. Stat Biosci. 2017;9:357–78.

105. Sweeting MJ, Thompson SG. Joint modelling of longitudinal and time-to-event data with application to predicting abdominal aortic aneurysm growth and rupture. Biom J. 2011;53:750–63.

106. Yuen HP, Mackinnon A. Performance of joint modelling of time-to-event data with time-dependent predictors: An assessment based on transition to psychosis data. PeerJ. 2016;2016.

107. Rizopoulos D, Hatfield LA, Carlin BP, Takkenberg JJM. Combining dynamic predictions from joint models for longitudinal and time-to-event data using bayesian model averaging. J Am Stat Assoc. 2014;109:1385–97.

108. Janes H, Pepe M, Gu W. Assessing the value of risk predictions by using risk stratification tables. Ann Intern Med. 2008;149:751–60.

109. Li RS, Huang XL, Cortes J. Quantile residual life regression with longitudinal biomarker measurements for dynamic prediction. Journal of the Royal Statistical Society Series C-Applied Statistics. 2016;65:755–73.

110. Kuwaki K, Nouso K, Kobayashi Y, Nakamura S, Ito YM, Iwadou S, et al. Prognostic model for hepatocellular carcinoma with time-dependent factors. Acta Med Okayama. 2011;65:11–20.

111. Palazon-Bru A, Carbayo-Herencia JA, Vigo MI, Gil-Guillen VF. A method to construct a points system to predict cardiovascular disease considering repeated measures of risk factors. PeerJ. 2016;2016 (2) (.

112. Hamoen M, Vergouwe Y, Wijga AH, Heymans MW, Jaddoe VWV, Twisk JWR, et al. Dynamic prediction of childhood high blood pressure in a population-based birth cohort: a model development study. BMJ Open. 2018;8:e023912.

113. Sweeting MJ. Using predictions from a joint model for longitudinal and survival data to inform the optimal time of intervention in an abdominal aortic aneurysm screening programme. Biom J. 2017;59:1247–60.

114. Yu MG, Taylor JMG, Sandler HM. Individual prediction in prostate cancer studies using a joint longitudinal survival-cure model. J Am Stat Assoc. 2008;103:178–87.

115. Zhang N, Chen H, Zou Y. A joint model of binary and longitudinal data with non-ignorable missingness, with application to marital stress and late-life major depression in women. J Appl Stat. 2014;41:1028–39.

116. Zhang S, Mueller P, Do K-A. A bayesian semiparametric survival model with longitudinal markers. Biometrics. 2010;66:435–43.

117. Duan LL, Wang X, Clancy JP, Szczesniak RD. Joint hierarchical Gaussian process model with application to personalized prediction in medical monitoring. Stat (Int Stat Inst). 2018;7.

118. Tseng YJ, Ping XO, Liang JD, Yang PM, Huang GT, Lai FP. Multiple-time-series clinical data processing for classification with merging algorithm and statistical measures. IEEE J Biomed Health Inform. 2015;19:1036–43.

119. Zhang D, Shen D, Alzheimer's disease neuroimaging I. Predicting future clinical changes of MCI patients using longitudinal and multimodal biomarkers. PLoS ONE [Electronic Resource]. 2012;7:e33182.

Publisher's Note