MDPI

Review

Clinical predictive modeling of heart failure: domain description, models' characteristics and literature review

Igor Odrobina

Mathematical Institute, Slovak Academy of Science, Štefánikova 49, SK-841 73 Bratislava, Slovakia; igor.odrobina@mat.savba.sk

Abstract: This study attempts to identify and briefly describe current directions in applied and theoretical clinical prediction research. Context-rich chronic heart failure syndrome (CHFS) telemedicine provides the medical foundation for this effort. In the chronic stage of heart failure, there are sudden exacerbations of syndromes with subsequent hospitalizations, which are called acute decompensation of heart failure (ADHF). These decompensations are the subject of diagnostic and prognostic predictions. The primary purpose of ADHF predictions is to clarify the current and future health status of patients and subsequently optimize therapeutic responses. We proposed a simpified discrete-state disease model as an attempt at a typical summarization of a medical subject before starting predictive modeling. The study tries also to structure the essential common characteristics of models in order to understand the issue in an application context. The last part provides an overview of prediction works in the field of CHFS. These three parts provide the reader with a comprehensive view of quantitative clinical predictive modeling in heart failure telemedicine with an emphasis on several key general aspects. The target audience is medical researchers seeking to align their clinical studies with prognostic or diagnostic predictive modeling, as well as other predictive researchers.

Keywords: prediction, model, heart failure, telemedicine, prognosis, diagnosis, detection, monitoring, characteristic

Citation: Odrobina, I. Clinical predictive modeling of heart failure: domain description, models' characteristics and literature review. *Diagnostics* **2023**, *1*, 0. https://doi.org/

Received:
Revised:
Accepted:
Published:

Copyright: © 2023 by the author. Submitted to *Diagnostics* for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Digital data is currently available in abundance in all healthcare facilities. Once automated analysis and prediction systems are built, researchers could realize a complete real-time analytical and decision support system. The widespread presence of electronic health records (EHRs) is also changing clinical prediction and analytical research. The new possibilities of data-driven research are pointed out in [1].

Modern applied and theoretical clinical prediction research bridges medicine, statistics, machine learning (ML) and engineering. All these areas have their own methods and terminology. A researcher trying to understand this field must become familiar with minimum basics in these fields.

In order to gain a representative sample of already applied predictive models, we focused on the well studied topic of telemedicine care for patients with chronic heart failure syndrome. We were driven by expectations that this approach, with some modifications, would form the framework for identifying cutting-edge topics and procedures in clinical predictive modeling. We found that a wide variety of statistical and machine learning models have been used in this area. In order to grasp the topic in its entirety, we have divided the study in accordance with its title into three parts.

In the first part, information of the disease was compiled primarily from medical journals. We aimed for a modeling systematized description of the CHFS disease. The compiled medical information is condensed into two diagrams in the UML style, which allows the consideration and optimization of the deployment of quantitative models. These two diagrams represent a simplified model consisting of individual disease states.

Understanding the general characteristics of quantitative models is critical to understanding the wide variety of published works. In addition, it allows the researcher to compare the predictive properties of the models and consider their possible new clinical applications. In the second part, we try to encompass the quantitative models into an abstract structure consisting of a single mathematical or algorithmic core surrounded by its external elements. The external elements we will call *common characteristics*. Partial classification of prediction models were made in [2,3] and general modeling recommendations were formulated in [4,5]. In this study, we do not deal with the methodology of model construction and deployment. We seek to develop a classification framework to identify at least some of the gaps in researchers' current prediction efforts.

We should also note that there are two fundamentally different prediction tasks in the investigated area. The first task is to quantitatively determine the prognosis of the patient's decompensation or death. The prognosis is then expressed as an individualized risk value over a relatively long period. The second task results from the temporal nature of telemedicine monitoring. The task is to continuously perform detection or diagnosis of the early stages of ADHF decompensation. The process can be used for a timely and optimal therapeutic response.

The third part of the study consists primarily of an overview of predictive modeling studies in the field of heart failure syndrome. After completing the review of publications on the prediction of CHFS, we expanded the literature search to include additional directions. First, it was the direction of *early warning systems* used in emergency rooms (ERs) and intensive care units (ICUs). The primary purpose of these systems is the *detection* of impending adverse health conditions. Other additional directions were related to the subsection dedicated to advanced statistical models to include other types of diseases. The last direction was related to subsection describing alternative modeling approaches. This was done in an attempt to understand the theoretical possibilities of the predictive ability of the models used in contemporary medicine.

1.1. Content and Structure of the Study

The study was organized as follows. In the *Medical Domain Description* section, we present an introduction to the clinical basis of chronic heart failure syndrome and the telemedicine of the syndrome. We describe the syndrome and its decompensation denoted as ADHF from a medical and modeling point of view. These subsections are followed by a description of telemedical patient remote monitoring. Section *Clinical Prediction Models in General* provides the reader with a general description and guidelines for medical prediction research. The section *Common Characteristics of Prediction Models* presents a classification of prediction models according to their outer characteristics. In it, we elaborate on topics such as the object of prediction, the form and timeline of predictive information, timelines of target and predictor data, and groups and types of prediction data. We discuss prognostic and diagnostic alternatives to model focus. At the end of the section, we discuss the machine learning approach in relation to traditional statistical modeling. Section *Mathematics of Quantitative Models: Two Simple Examples* presents mathematical formulations of introductory diagnostic and prognostic prediction tasks. *Overview of Heart Failure Prediction Models* first provides an overview of remote monitoring with attempts to perform diagnostic or detection prediction of decompensations. Next, we provide a link to several review studies that prognostically predict patient decompensations. This is followed by a subsection describing advanced statistical models and other statistical models. At the end of the section, we summarize prognostic and diagnostic predictions carried out by machine learning researchers. *Summary and Conclusion* provides a closure of this study.

1.2. Literature Search Method

The publication database MEDLINE represents an overwhelming variety of medical topics and directions. In our opinion, no single review study can fully satisfy the reader interested in a particular field.

In order to meet the objectives formulated in the abstract, we tried to conduct our searches as cross-sectional as possible. Our literature review selects only a few publications from a particular research direction that we hope are the most recent or the most comprehensive. We are not trying to do a complete review of any specific direction. The most numerous articles on predictive modeling are those in which the modeling is based on data present in the EHR. They usually predict ADHF decompensation and hospital readmission as a prognosis for CHFS. These prognostic predictions are mainly based on biomarkers, one-time medical examinations and demographic data.

Not many reviews have been published on the topic of detection or predictive diagnosis of ADHF in the telemedical settings. This prediction is based on data that is collected with a relatively high repetitive rate and the data constitutes a special kind of EHRs.

In the field of advanced statistical models, we found only a few articles dealing with the prognosis of CHFS yet, but we present these models for a more complete review. We included two publications dealing with the prognostic prediction of CHFS incidence or *incident heart failure*.

Our review is based on searches performed primarily in the MEDLINE database via the PubMed search engine. The results were confronted with the findings in Google Scholar. We consider the scope of the topic under investigation to be so diverse that we have not attempted to cover our findings within some unifying search query and search scheme. Obviously, in the early search stages we created a complex search query that included words like *prediction*, *model*, *heart failure*, *telemedicine* or *remote monitoring*, but later we focused on other search methods to identify desired publications, such as snowballing through article links.

2. Medical Domain Description

2.1. Chronic Heart Failure Syndrome and its Decompensation

Chronic heart failure syndrome is a frequent and long-term disease that burdens the patient's life and represents also a burden on the medical care system. The syndrome arises as a result of various worsened underlying cardiovascular health conditions. In general, after the appearance of typical signs and symptoms, comprehensive examinations of CHFS are performed before the diagnosis is finally confirmed [6,7]. CHFS is often grouped into two categories according to the status of left ventricular ejection fraction (LVEF). Heart failure in patients with reduced LVEF is referred to as HFrEF, in patients with preserved LVEF as HFpEF. Patients are further classified according to the NYHA scale and the score of the KCCQ-12 questionnaire. Many patients progress to a stage called *advanced heart failure* characterized by persistent symptoms [7].

In addition to the slow continuous deterioration of the quality of life, the patient's life is disrupted by a sudden worsening of symptoms, which is called *acute decompensation of heart failure* or ADHF. For simplicity, in this work we will adopt this terminology used in [6]. In European guidelines [7] there is a term *acute heart failure* or AHF and acute decompensation is referred to as only one of four different types of acute presentations. In addition to acute decompensation, other three presentations are acute pulmonary edema, isolated right ventricular failure, and cardiogenic shock.

We could now say that our broader definition of ADHF now includes four distinct presentations [8] with different temporal characteristics of progression. We will discuss the implications of these issues in the modeler's subsection. The differences in terminology mentioned above were addressed in [9].

At the end of this section, we could mention that the worsening of symptoms requiring hospitalization as the beginning of CHFS is called *de-novo acute heart failure* [8,10]. It is separate topic and we do not deal with it here. If one wants to understand the extent of the medical and biochemical models of CHFS, one should look at the works of D. L. Mann et al. [11].

2.2. Description of Heart Failure Syndrome from a Modelers' Perspective

In addition to biochemical modeling, two additional modeling processes appear have to be run in parallel in CHFS prediction task. The first additional modeling process tries to grasp complicated medical diagnostic-prognostic contexts. The person performing this activity should be called a medical *domain expert* and should describe the investigated problem in some form of modeling language such as UML. The second modeling process refers to quantitative predictive modeling, where a mutual combination of predictors and a mathematical formula explain the predicted parameter. The person performing this modeling process is a statistical or machine learning expert. In the study, we do not always distinguish between these two cases.

The modelers do not necessarily need to know every detail of the biochemistry of the investigated health condition, but they need to know the basics about timelines, predictors, manifestations and all possible outcomes of the disease. They should also be aware of the fact that ADHF decompensation is a relatively autonomous biomedical pathological sub-process of CHFS with a more or less well-defined onset and end.

The health information about CHFS in the previous paragraphs can be summarized in the form of a four-state diagram in the figure 1. To the basic three states (chronic state, acute decompensation state and terminal stage) we added a fourth state - the advanced state in accordance with [7]. According to [7, Chapter 4], the weight of the patient in the compensated and advanced state develops in the opposite direction, which may indicate primarily different disease states not only from a diagnostic, but also a prognostic point of view. A change in the patient's body weight is a key sign of heart failure syndrome.

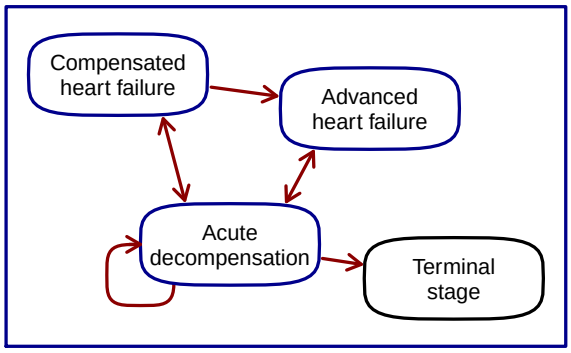

Figure 1. Diagram of states and transitions of chronic heart failure syndrome. Red straight arrow indicates state transition. Curved arrow indicates a state recurrence. The terminal stage refers to the patient's irreversible progression to death.

This is the disease modeling approach advocated by Houwelingen [12] and others. In their model formulation, transitions are associated with rates or probabilities. These models are primarily suitable for prognostic quantitative predictions. It should be noted that the acute decompensation state does not have the so-called *memoryless* property. This is related to the fact that the ADHF state is preferentially restored to the state from which it originated. ADHF can also recur, and then the patient is re-hospitalized within a month or two of discharge.

The state diagram in Figure 1 can be unfolded into a temporal progression of CHFS states, as shown in Figure 2. When the syndrome states are arranged according to disease severity on the vertical axis, a striking correspodence with Figure 1 in [13] appears. Although the dark blue line in our model represents the patient's time-varying states, the diagram in Figure 2 resembles some discretization of the patient's clinical status development in Figure 1 in [13]. The similarity is even stronger when we consider that the *Advanced*

Heart Failure state corresponds in meaning to the *Chronic Decompensated* label on the left axis. We find this remarkable and this alternative model type is discussed below.

In Figure 2, the first transition from the compensated state to the decompensated ADHF state is explicitly marked. Towards the end of the diagram ADHF recurrence is illustrated.

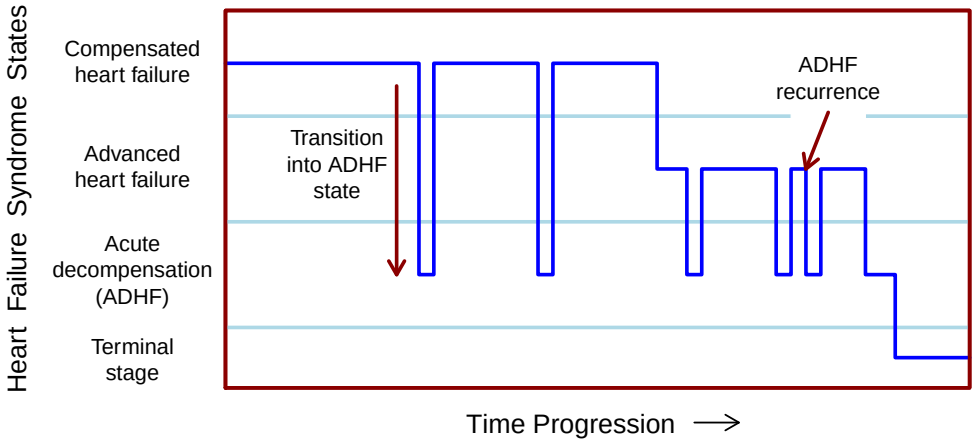

Figure 2. Illustration of the progression of chronic heart failure syndrome as a sequence of time-varying states. The light blue horizontal lines mark the boundaries of the four syndrome states. Dark blue line indicates patient's state. Transitions between states are depicted as instantaneous.

The Figures 1 and 2 together can represent a working model of the clinical states and events in a patient with chronic heart failure syndrome. It can be considered as a prognostic discrete-state version of the CHFS model suggested in [13]. All prognostic studies reviewed in section *Overview of Heart Failure Prediction Models* are based on simplifications of this model into two-state forms. Compensated and advanced heart failure states are merged together; and published studies focus separately on transitions to acute decompensation and transitions to the terminal stage. In the latter case, the ADHF state is merged to the terminal stage.

The chronic heart failure model suggested in Figure 1 in [13] is of a different type. The model is highly predictive and could probably be developed so that clinical status represents a disease indicator carefully designed as a combination of patient diagnostic parameters. As we will show later, this type of model is able to clarify the telemedicine diagnostic processes of CHFS.

It would also be useful to know whether ADHF is triggered by some random cause (external or internal), or whether decompensation occurs as a natural internal progression of CHFS. A list of probable random causes triggering ADHF decompensation and their statistics are given in [14][1].

It is also important to have unambiguous specifications of the outcome events. The basic adverse outcome event can be an irregular visit to the ambulance, hospitalization or even death. Each type of event can have its own optimized set of predictor variables. As previously mentioned, a patient admitted to the hospital with our more broadly defined ADHF may have four different clinical presentations. Prognostic prediction of new-onset heart failure syndrome [15] can serve as an example of the substantial impact of detailed outcome specification on a set of optimized predictors.

The modelers should understand the underlying dynamics of the predicted acute process. There can be several types of events, and the onset of the event can be gradual (days) or rapid (hours) or indeterminate [7, Chapter 11]. They should have an idea of

[1] We should note that we have not found much independent support for these observations in literature yet.

the nature of the symptoms and be aware of main clinical manifestations. In the case of a controlled trial, this is important for the selection of the correct set of measured medical parameters. In the case of a retrospective observational study, clinical information is important at least to eliminate the presence of outliers and systemic outliers.

At this point, we should clarify the situation with insufficiently clear boundaries between key medical concepts. In the following text, we will simply assume that there is no transition period between the compensated stage of CHFS and the acute decompensated stage (ADHF). This is in apparent contrast to the designation expressed in the title of the publication [16]. We will consider the transition period forementioned in the title as some early stages of acute decompensation process. These early stages are manifested, for example, by changes in patient's pulmonary arterial pressure.

At the end of the subsection, we mention that the understanding of important concepts in the field of CHFS is hindered by the fact that the global medical community follows two rather different systems of reasoning, characterized by two separate guidelines [6,7]. The modelers should also be aware that the syndrom is characterized by non-specific symptoms and signs [17], that there is no single test to establish the diagnosis of CHFS [18] and that 14 - 29% of cases are misclassified even after examination in emergency room [19].

2.3. Telemedical Remote Monitoring of Patients with Heart Failure

Modern telecommunication technologies have also penetrated the field of health care for patients with chronic heart failure syndrome. These technologies make it possible to use a hitherto unused set of data describing the patient's signs and symptoms, which are collected during the ordinary life of the patient on a daily basis or even more often. This data has remained unused until now despite its importance [20]. The importance of collecting this type of data in the home environment is also documented by the CHFS guidelines [7], which says, for example, that if the patient's weight increases above a certain level over a certain period, the therapist or the patient himself should administer an increased dose of diuretics.

Medical staff in telemedicine trials now have unrestricted access to this daily data in parallel with the patient's biomarkers and medical examinations obtained during initial or regular visits. The therapist now has the opportunity to use them to adjust his action in order to ensure the best long-term prognosis for the patient. It should be noted that the primary role of remote patient monitoring in CHFS telemedicine is to *improve patient medical management*; decompensation prediction is only a subset of this primary assignment. The authors [21] hypothesize that *the most potent therapeutic effect of telemedicine comes from more optimal use of diuretics and up-titration of additional CHFS medication*. Optimization of medication doses based on CHFS telemedicine data was investigated in [22].

As can be seen in the review by [23], before 2002, telemedicine data was collected in a non-invasive way, i.e. without any wearable devices and implants. But currently, the number of projects using invasive methods of remote monitoring of patients with CHFS is growing. Recently, review articles [24–26] attempted to evaluate the overall impact of implant-based telemonitoring on the management of patients with CHFS. It should be noted that distrustful views have also been expressed about this technology [27]. Wearable devices in this context have been investigated in [28].

On the other hand, in addition to invasive and device-assisted methods, there are still many new non-invasive telemonitoring studies in chronic heart failure medicine. A survey [29] found that CHFS telemonitoring was associated with a 20% reduction in all-cause mortality and a 37% reduction in CHFS hospitalization. Other CHFS telemedicine trials were reviewed in [30–33]. We could consider the work of [34] as the most promising study of non-invasive telemonitoring, which shows the positive benefits of telemedicine care above a statistically significant level. It is not self-evident that this outcome can be achieved, and many other telemedicine studies [35–38] show that CHFS telemedicine improves patient outcomes, but not as much as required by the 5% level of statistical significance.

The common characteristic of telemetry data is the relative simplicity of their monitoring, their collection is often done by the patient himself. The term *vital signs* usually refers to heart rate, blood pressure, respiratory rate, and body temperature, but we prefer to use the term more loosely as a category of data collected at a high repetition rate that also includes weight change, oxygen saturation level, and the onset or worsening of symptoms. We will return to this issue later in the discussion of grouping data types.

3. Clinical Prediction Models in General

Clinical prediction and clinical prediction tools are an integral part of modern medicine. A large number of prediction models are published every year. The basics of predictive modeling in medicine are summarized in [5]. Chapters aimed at a medical audience have been included, such as *Predictive Modeling Studies*, *Predictive Model Applications*, and more. Systematic evaluation of the clinical utility of predictive modeling is a complex task and requires a *decision* and *analytical* framework [39]. Another team of authors evaluated the impact of prediction models in [40]. More work of this kind is needed to clarify the medical foundations of prediction research and to overcome the doubts that have been directed at it like in [41].

Given the diversity and complexity of the prediction research community and prediction research itself, there are also efforts to guide the research and reporting process by specifying a fixed set of rules. Intuitive and disorganized reporting of developed models can very easily devalue the primary achievements and messages of the authors. Therefore, a joint effort to structure and regulate the issue of model reporting appeared. An initiative called *Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis* or the TRIPOD [42,43] came into existence, in which the basic principles are explicitly formulated. Methodological guidance for models' updating can be found in [39].

It is well known that models are often subject to bias. Another initiative emerged and developed the *Prediction model Risk Of Bias Assessment Tool* or PROBAST tool [44,45]. The tool consists of four fields: participants, predictors, outcome and analysis. These domains contain a total of twenty *signaling questions* to assess risk of bias. The level of risk of bias generally depends on the study design, conduct and analysis. A high risk of bias indicates a significantly distorted performance of the model's predictive output.

Very valuable information about predictive modeling and the properties of statistical models can also be obtained through area-specific guidelines [46] and systematic reviews [47]. A practical guide to clinical prediction modeling can be found in [2].

4. Common Characteristics of Prediction Models

The goal of this part is to provide a structured, unified view of quantitative predictive models in statistics, engineering, and machine learning. These fields attempt to solve the prediction task defined in the medical domain of chronic heart failure syndrome, where the situation is captured by the disease model condensed in Figures 1 and 2. We consider the introduction of this view as an analogy to the introduction of an additional type of diagrams when describing an domain problem in UML language.

As we have already mentioned, prediction models can be assessed according to their external elements or characteristics. External model elements or characteristics can be introduced as features of the model that do not belong to the internal statistical or algorithmic core. They represent a kind of surroundings of the model core interior. They are shown in lower part of the Figure 3. Of all the characteristics present in the model, in this part we focus on the object of prediction, the time characteristic of diagnostic and prognostic information, target and predictor data, and types and groups of prediction data. The set of model characteristics also includes information on whether the model deals with prognostics or diagnostics and whether a statistical core or a machine learning core was used. At the end of this section, we present two simple examples of mathematical model cores.

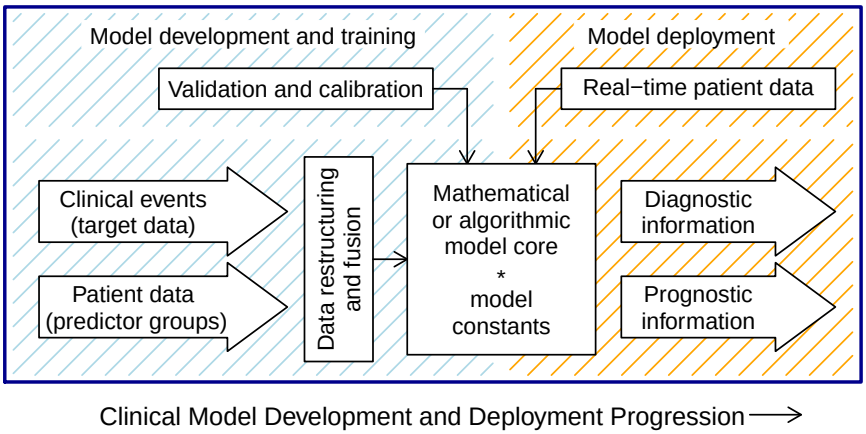

Figure 3. Schematic representation of the clinical quantitative modeling and model deployment.

Before continuing, we would like to remind non-mathematicians that the concept of probability or risk of developing a disease can be imagined as the proportion of materialized positive cases within a relevant cohort of patients in a time interval. The next discussed *hazard rate* can then be understood as this probability divided by the mentioned time interval.

4.1. Characteristic #1. Object of Prediction

A prediction is a statement about a clinically relevant issue that is in a state of uncertainty at the moment of prediction. The concept of mathematical probability is used to quantitatively express prediction. In clinical practice, there is uncertainty about the presence of the disease or its stage at the moment of prediction. A prediction can also be a quantitative probability statement about the occurrence of a disease or its stage in the future.

In the context of already diagnosed CHFS, the focus of prognostic and diagnostic predictions shifts to the occurrence of worsening symptoms, the appearance of a stage of decompensation with admission to the hospital, or the occurence of death. These are all visible manifestations of a sudden change in the compensated CHFS state. The primary aim of this study is to investigate the prognostic and diagnostic prediction possibilities of ADHF in patients with an already established diagnosis of CHFS who are under telemedicine monitoring. The model review part also includes works with prognostic predictions of deaths.

4.2. Characteristic #2. Prediction Information Timelines

4.2.1. Diagnostic Information Timelines

Prediction can be aimed at predicting the presence of a disease or its stage in a patient at the current moment. This is a diagnostic prediction. The meaning of the word prediction seems to be related primarily to the uncertain nature of the prediction statement. Diagnostic uncertainty fades over time in two ways. The first is related to the timeline of disease progression, when the disease manifests itself with more intense and visible symptoms. The second is connected to the timeline of the sequence of diagnostic steps, when more accurate and unambiguous tests are applied later in the sequence.

In the first case, the signs and symptoms of the disease or its new stage are detected and the diagnosis is predicted. The validity of the prediction is confirmed by the explicit manifestation of the disease only with a certain time delay, which is clearly shorter than the duration of the entire pathological process, which in our case is ADHF. Diagnosis or detection of ADHF by measurement of pulmonary arterial pressure may precede hospital admission by approximately twenty days [16]. The certainty of diagnostic prediction is

quantitatively expressed by the values of sensitivity and specificity,[2] when larger values mean greater certainty. The quality of the entire diagnostic method is assessed by the Area Under the Receiver Operating Characteristic (AUROC) curve.

In the field of CHFS telemedicine, the diagnostic procedure is carried out remotely regularly with a high frequency of repetition. The moment of diagnosis moves forward, and as hospitalization approaches, the certainty of diagnosis should change towards higher values.

In the second case, in the case of a sequence of diagnostic tests, the prediction is refined by applying more accurate additional tests. The therapist makes a decision about the disease not only on the basis of a more accurate test, but also considers the results of the previous ones. The issue of combining information from several diagnostic tests or symptoms is of fundamental importance, and its mathematical description is discussed in section *Mathematics of Quantitative Models: Two Simple Examples*.

4.2.2. Prognostic Information Timelines

The presence of ADHF is manifested by the event of the patient's admission to the hospital. The decompensations are said to occur randomly, so their manifestations.

There are two distinct types of prognostic prediction of decompensation in the literature. The first type of prognostic prediction is the prediction of the occurrence of decompensation in the near and distant future. The second type concerns only the near future, which means that the time interval for the rate or probability calculation starts from the moment of prognostic prediction.

During the modeling process, the *prognostic period* corresponds to the entire period of the follow-up study. The prognostic period should be much longer than the typical duration of decompensation.

For the first type of prognostic prediction, the powerful concept of *hazard rate function* [48–50] is widely used. The hazard rate or frequency of decompensations in a patient cohort may change over a relatively long prognostic period. This is why the time-dependent function is used to capture the prognostic information as a whole.

A precise definition of the hazard rate function can be made through its relation to the probability of an event or probability of change in the disease state denoted P. Mathematically, it can be expressed as follows. First, the randomness of a disease event is described by a random variable T which represents the time of occurrence of the event. The hazard rate function $h(t)$ is then defined as the rate of occurrence of events at time t. Time t is positive and less than or equal to the prognostic period. Using the formalism of probability equations, this can be expressed as [48–50]:

$$h(t) \simeq \frac{P(t \leq T < T + \Delta t | T \geq t)}{\Delta t} \,, \tag{1}$$

where Δt is the interval for counting events to obtain an observational estimate of the probability P and should be long enough to eliminate statistical noise. The Δt is not directly related to the duration of the decompensation process (ADHF) but must be reasonably longer than the duration of its manifestation (e.g. the duration of the hospital admission acceptation process). It is usually much shorter than the prognostic period.

In the context of CHFS prognosis, the expression (1) reads that the hazard rate $h(t)$ is the rate at which patients in the cohort experience the occurrence of decompensation. The condition $T \geq t$ in the conditional probability says that the calculation of the proportion takes into account only those patients who have not experienced the event until time t.

The hazard rate function can be constant, increasing, U-shaped, or shaped in some other way, as shown for example in [51]. To get a sense of the possible statistical noise distortion of the observed hazard rate functions, one should look at the examples in [48, Chapter 2]. The Kaplan-Meyer, Nelson-Aalen, and Cox model with its variants are used to

[2] Quantitative definitions of sensitivity and specificity are given in Table A2.

calculate the hazard rate function. The hazard rate value obtained for the whole cohort can be individualized according to individual patient characteristics, as discussed in section *Mathematics of Quantitative Models: Two Simple Examples*. The quality of the prognosis can be evaluated using time-dependent receiver operating characteristic (ROC) curves [52].

The second simplistic type of prognostic prediction is a prediction for the near or impending future. The prognostic period corresponds to the counting interval Δt and a constant value of the hazard rate is assumed. The interval can be as long as a day, a week, a month, a year, even as long as the patient's remaining life. The outcome events are counted together during the entire follow-up period. In the context of CHFS, the number of outcome events represents the *cumulative incidence* of decompensations. During this period, the group of patients is partially reduced, but the period can be chosen short enough not to significantly affect the modeler's quantitative predictions. The cumulative probability of events p is calculated in the interval Δt and has the form of a simple equation:

$$p(\Delta t) = P(0 \leq T < \Delta t | T \geq 0) \ ,$$

where T is again a random variable assigned to the time of the event. When the time interval Δt is reasonably short, the relationship between the cumulative probability p and the hazard rate function $h(t)$ can be expressed using approximate equality:

$$p(\Delta t) \simeq h(t)|_{t=0} \, \Delta t \ .$$

The approximate equality can provide a quick estimate of the hazard rate when the proportion of the cumulative incidence of events is less than some 10-20% of the total number of patients. The well-established logistic regression is widely used in this type of prognostic prediction.

4.3. Characteristic #3. Temporal Properties of Target and Predictor Data

4.3.1. Temporal Properties of Target Data

Target data represent basic information about recorded clinical events. In a simple diagnostic prediction model, the data need not to have explicit temporal characteristics. If continual diagnosis of monitored patient is performed, the target data can be bounded by a sliding time window that moves with the moment of diagnosis.

In case we are building a model specifying the prognostic hazard rate function, we need to have event time data in the data set. They are present there in the form e.g. that a patient event record contains the patient ID, event time, and event type coded into a categorical variable. The role of this target *time-to-event* data in models is significantly different from the role of time data specifying the time of the predictor value. A simpler, previously defined second type of prognostic prediction model does not require the precise specification of the time of the event. The length of the follow-up period, which is equal to the length of the counting interval Δt, is sufficient.

4.3.2. Temporal Properties of Predictor Data

Incorporating time dependence into predictor variables seems to be one of the primary challenges of prediction models in contemporary prediction research. We call statistical models that directly include the time dependence of predictors *advanced models* due to a significant increase in their complexity.

First, in the simplest case, the predictor variables have no significant time dependence at all. Predictor data is collected over a time period of negligible length. In the context of CHFS research, this is the case of a patient's entry into a clinical trial or case of a hospital entry examination to confirm ADHF diagnosis. Over the course of the clinical trial, data is not updated, and information about the patient's ever-changing vital signs and symptoms is not collected or ignored. These time-free data represent pure *cross-sectional* data.

The second case occurs when a patient visits a therapist during a clinical trial and their biomarker and other data are updated on a quarterly or monthly basis. This data usually

contains a time dependence, but the data update frequency is relatively low. This *episodic or regular visit data* enters the prediction models in a significantly different way than the target time-to-event data. This data is called *longitudinal data*. Models using this type of predictor data are summarized in the subsections *Advanced Models* and *Other Models*.

The third case occurs when vital signs and disease symptoms are recorded and actively incorporated into the modeling in a telemedicine clinical trial. These types of data are collected at a significantly higher repetition rate compared to the previous case. In telemedicine or home patient care settings, this data is collected daily or almost daily. For intensive care unit vital signs, the collection rate can be hours or even minutes. The term longitudinal data is very rarely used in the literature for these data, and the term *time series* seems to be preferred. In the context of telemedicine, these data are used in predictive models to detect or, more precisely, to diagnose the onset of acute decompensation of heart failure. In the intensive care unit, this data is used to probabilistically determine e.g. the 24-hour risk of adverse events such as cardiac arrest.

4.3.3. Temporal Properties of Input Data during Continuous Diagnosis

During the determination of the sensitivity and specificity of the method of diagnosis, the target input data represent the delayed explicit manifestations of the disease. The determination requires a certain time interval to compile target data to confirm or disprove the validity of the disease prediction. We can call the chosen time interval *forward target window*, and it should be large enough to cover the mentioned manifestation delays. In the field of CHFS telemedicine, the essence of patient monitoring is a process of continuous repetitive diagnosis. The moment of diagnosis is constantly shifting in time, as is the beginning of the forward target window.

In diagnostic prediction modeling, another time interval appears. The interval includes the temporal changes and temporal patterns present in the recorded predictor data. When monitoring a patient continuously, it is easy to include multiple records from the recent past. They could also serve to eliminate random noise from recordings. These past data are again part of the diagnostic prediction process and can be considered as part of another time window, which we could call the *retrospective predictor window*. This window also moves with the progress of the diagnostic moment. The two prediction windows mentioned above could together be termed as *sliding time windows* [53]. The introduction of similar windows is also present in other works and in the field of continuous diagnosis it represents an additional form of input data structuring.

The relation of both windows to the development of diagnostic parameters is shown in Figure 4. The schematic describes a retrospective modeling situation, so we know with certainty that in this case the hospitalization event definitely occurred. We can rescale the time axis so that the moment of hospitalization correspond to time zero. The retrospective predictor window specifies the range for the predictor data, which are plotted in the figure by the blue line. The forward target window determines the range of the target data. The only target value in the scheme is represented by the act of hospitalization at time zero. The moment of diagnosis is marked with a dark red arrow. Figure 4 represents a more general view of continuous detection-diagnostic prediction process investigated e.g. in [54].

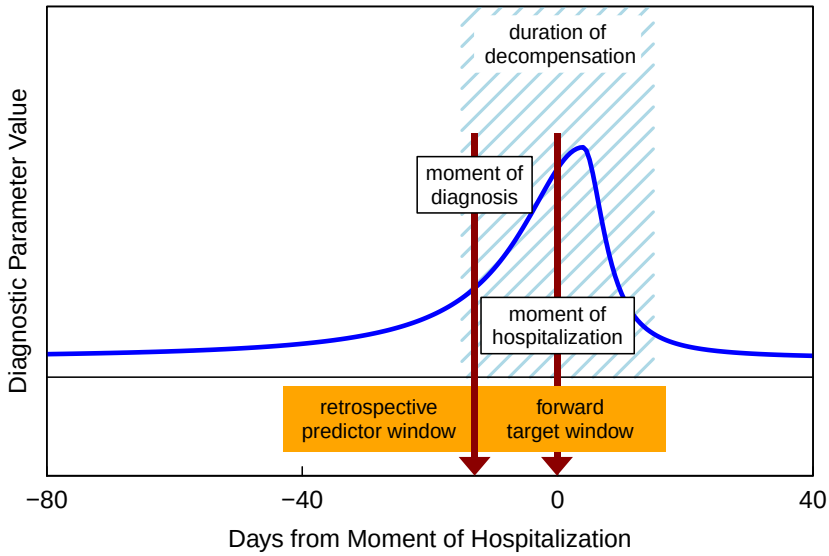

Figure 4. Schematic of continuous telemedical diagnostics against the background of patient decompensation (hatched area). The illustration of the diagnostic parameter development (blue line) is made in accordance with [55,56].

Provided that the clinical status parameter in suggested heart failure model in Figure 1 in [13] is linked with the patient's diagnostic parameters, this model has the ability to clarify the diagnostic processes of acute decompensation (ADHF) during telemedicine monitoring. The overall picture of the telemedicine continuous diagnosis of heart failure patient can be obtained by gradual superimposing the inverted diagnostic parameter curve in Figure 4 (blue line) over the pronounced depressions on the patient's clinical status curve in [13].

4.4. Characteristic #4. Processing of Different Types and Groups of Predictor Data

The timelines for a typical telemedicine controlled trial are as follows. The telemedicine study begins with an entrance examination of both the control and intervention groups. The study continues with telemedicine monitoring of the intervention group, which may last half a year or longer. Telemedicine data of the intervention group are collected in the home environment daily or almost daily. Data on regular and occasional visits to the ambulance are also stored. At the end of the clinical trial, both groups will undergo a final exit examination and the results will be used for comparison. According to the temporal characteristics, the data can be classified into five groups, as shown in the table 1.

Table 1. Different data groups in a typical telemedical CHFS trial.

Data group	Temporal characteristics
Demographic data (baseline data)	Not changing in time
Entry examination data (baseline data)	Collected at the time of entry examination
Telemedicine data (vital signs*)	Time dependent (high repetitive rate)
Episodical or regular visit data	Time dependent (low repetitive rate or episodic)
Final examination data	Collected at the time of final examination

* By *vital signs* we denote all easily obtainable patient characteristics measured with high repetitive rate.

A fundamental aspect of these data is their heterogeneity in relation to time. As we can see, some types of data are collected only once or twice, another group of data is collected episodically or with a low frequency (monthly, quarterly), and some data is collected with a high repetition frequency (daily).

There are obvious similarities between our ADHF telemedicine diagnostic system and well-established early warning systems used in ERs and ICUs. Heterogeneous data types are present in these systems too, and we consider the prediction experience accumulated in this area to be substantial also in telemedicine field. The combination of several different groups of predictors with different temporal characteristics into one prediction process has been labeled as a *data fusion* [1, Chapter 22]. The authors build on predictive modeling works [57,58] and sort their data into groups in the manner shown in the Table 2. By developing the presentation of Table 2 we kept the original structuring of the data, but the medical parameters were changed to correspond to our CHFS area. The presented sets of CHFS trial data types are for demonstrative purposes, they are compiled from literature and project proposals by non-medical expert and should be considered illustrative only.

Table 2. Data types structured into groups for the data fusion method.

Entry examination (baseline data)	**Vital signs* (highly repetitive data)**
NYHA II - IV	Heart rate
LVEF	Systolic blood pressure
ECG	Diastolic blood pressure
Haemoglobin	Body weight
Serum sodium	Oxygen saturation
Serum potassium	Symptom intensity level
Serum creatine	
NT-proBNP	**Demographics (baseline data)**
CRP	Age
BUN	Race
KCCQ-12	Gender
6-minute walk test	

* By *vital signs* we denote all easily obtainable patient characteristics measured with high repetitive rate.

Comparing Table 2 with Table 1, we see that the final examination data group and the low repetition frequency data group are missing. Nevertheless, we believe that the ideas of the *data fusion* technique are also applicable to telemedicine trials. A similar approach to the *data fusion* technique can be found in [59] for the early diagnosis and detection of ADHF in a telemedicine settings.

For a better prediction success rate, we could compose new predictor variables that could contain information about the time derivative or time integral characteristics of the originally observed predictors [60]. Basic statistics literature [48, Chapter 8] recommends creating these new variables as well. As an example, the relative time derivative of the observed variable is created. This new variable served to capture time dependence in prediction variables in a standard Cox proportional hazards model.

The term *feature engineering* is used in both the engineering and machine learning communities for the process of creating new, directly unobserved variables. Publications [61, 62] present a list of engineered variables from telemedically monitored daily data of patients with CHFS in order to diagnose ADHF.

Monitored telemedicine prediction data can be processed to create a predictive alert signal. In [54,63], an extended moving average method called MACD is used to generate a warning signal from a single monitored variable such as body weight. The *pattern similarity principle* is used to generate an alert signal from monitored patient vital signs in [53]. The predictive ability of individual signals can be strengthened by combining them with each other using the naive Bayesian assumption [61,62,64].

4.5. Characteristic #5. Distinction between Prognosis and Diagnosis

A natural start to understanding the distinction between the terms of prognosis and diagnosis is to follow the timelines of these predictions. Prognosis deals with the situation

where the pathological process of the disease is predicted to occur sometime in the future. Diagnostics deals with the current situation and assesses whether the disease process has started or not. We could repeat the statement in [39] that *clinical prediction models are tools that predict health outcomes either in the present (diagnostic) or in the future (prognostic)*. The difference between these two prediction categories is also described in [42,65].

In more complex situations such as continuous patient monitoring of ADHF, the above distinction is inconvenient to clarify the situation. We prefer to use differences in clinical parameters. In the classification schematic [65], the authors underscored the observation of the presence of disease signs and symptoms as the predominant difference between diagnosis and prognosis. In the case of diagnosis, we can rely on the presence of signs and symptoms of the disease, while in the case of prognostic prediction, we could not do this, because the patient does not yet have the predicted state of health. For prognostic prediction, we should rely only on other patient clinical parameters, such as biomarker values, clinical examination results, etc.

One should not be confused by applying schematics [65] for prediction clasification in our CHFS field. In the schematics, the authors use the term *cross-sectional* to describe the process of diagnosis. But the authors use the term to describe the simultaneity between the moment of the latest prediction data and the moment of predicted state of health. This may cause some confusion because the term *cross-sectional* is often associated with predictor variables, and this type of variable is regularly used in prognosis.

It should be noted that in situations when it comes to a disease with a long and complex medical history, we are dealing with a relatively long sequence of prognoses and diagnoses. A diagram of the diagnostic-prognostic sequence undergone by a patient in acute decompensated stage is shown in Figure 5. We see that at the beginning, an impaired cardiovascular condition occurred and was diagnosed. Within the prognosis of impaired cardiovascular condition, there is a possibility that chronic heart failure syndrome may occur. Once the CHFS occurs and is diagnosed, the prognosis of the syndrome is that a decompensated state of ADHF may follow. In telemedicine monitoring, ADHF is pre-diagnosed (or detected) in an outpatient setting, followed by a confirmatory diagnosis of ADHF in a medical facility. Again, the ADHF state has its prognoses, such as recovery to a compensated state, recovery to a chronically decompensated state, readmission relapse, and unfortunately, death.

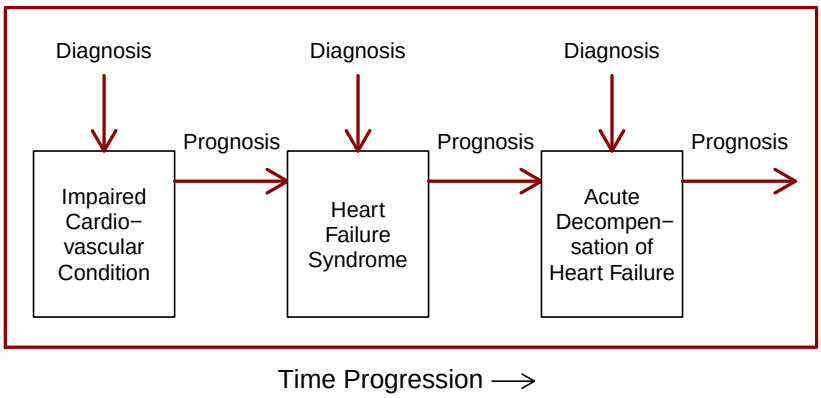

Figure 5. History of diagnoses and prognoses in a patient with ADHF.

Another term, *the detection*, is associated in the literature with the act of predicting a medical condition. It may come from the authors' engineering background as a convenient substitute for the term diagnosis [53,62,64]. However, the term *detection* also seems to be used in situations where the use of the term diagnosis is not easily applicable. This seems to be the case with early warning systems [66,67].

In this study, when predicting impending acute decompensation, we prefer the term *ADHF diagnosis* and follow the use of the term e.g. in [68]. We prefer to comply with the recommendations formulated in TRIPOD [42] and in [65].

4.6. Characteristic #6. Statistical Approach versus Machine Learning

The discussion on the relationship between the statistical approach and the machine learning was started by L. Breiman's article with valuable comments that express the position of several recognized statisticians [69]. Statistical approaches are based on a solid theoretical data model and the idea of likelihood in the background. Statistics also has its imperfect models, they call them *working models*. Mathematics is also part of machine learning. One must admit that there is a certain similarity between the search for maximum likelihood in statistics and the minimization of the error function in the field of neural networks. However, machine learning seems to be trying to build a perfect algorithm that provides perfect responses in response to input data, rather than building a perfect data model in the background.

In the area of prognostic survival modeling, these two approaches have been summarized [3] in understandable model hierarchies. Recently, machine learning has attracted critical attention from researchers with a medical background. In the area of ADHF prediction, a critical appraisal of ML studies was presented in [70]. Studies [71,72] address the issue of performance and reliability of machine learning models within a broader clinical perspective.

Despite the criticism, it should be emphasized that machine learning modeling represents a fundamentally different approach by researchers from different backgrounds. In theory, their challenges could prevent clinical prediction research from stiffness in methodology and concepts. On the other hand, it seems that machine learning researchers do not pay due attention to the fundamentals of clinical prediction formulated in the TRIPOD statement and PROBAST tool requirements.

Machine learning researchers have brought new concepts to the field of clinical prediction research that are not well established in statistics. Some of it comes from their engineering and especially their software engineering background. The term *domain expertise* draws attention to the fact that there is a relatively large area between modeling and medicine that does not belong to either modelers or medical professionals. Another term *conceptual embedding* describes the process of mapping clinical terminology to universal modeling concepts. In our opinion, this assignment deserves a clear name. Concepts used by healthcare professionals are formed by clinical practice and require additional specification before being used in quantitative modeling. Probably the most intriguing is the introduction of the concept *feature engineering*, which expresses the fact that modelers are not limited by the form of observed data and are encouraged to use their modifications as predictive variables. Machine learning experts introduced these concepts, probably because the application domain in their field often changes and this requires persistent flexibility.

4.7. Mathematics of Quantitative Models: Two Simple Examples[3]

Mathematics is present in all prediction models and plays a key role in model formulation and application. In the following text, we present a mathematical approach to two main problems in clinical prediction research.

4.7.1. Basic Diagnostic Model

The widespread availability of electronic health records makes it easy to conduct quantitative research on diagnostic procedures. A review of published sensitivities and specificities of symptoms for the diagnosis of ADHF was conducted in [19]. As we men-

[3] Reading this section is not necessary to understand the other sections.

tioned earlier, ADHF is characterized by symptoms of low specificity, and therefore the issue of combining diagnostic information from more than one symptom or sign is important.

From a mathematically exact point of view, the therapist performs a set of diagnostic steps where the initial posterior probability of the presence of the disease is constantly replaced by new improved posterior probabilities under new evidence. The mathematical explanation and formulation of the problem is as follows. The probability $P_1(D|E_1)$ of the presence of a disease state D at the result of the first diagnostic step E_1 can be expressed by the Bayes' theorem in perhaps the most transparent form as:

$$P_1(D|E_1) = \frac{P(E_1|D)}{P(E_1|D)P_0(D) + P(E_1|\neg D)P_0(\neg D)}\, P_0(D)\,, \tag{2}$$

where $P_1(D|E_1)$ represents the posterior probability of the disease, the expression $P_0(D)$ is the probability of the disease state D in the population. In case D indicates the presence of the disease, the expression above represents the prevalence of the disease (for further explanation see Table A2). The term $P(E_1|D)$ is the probability of the test result E_1 on the disease state D, $P(E_1|\neg D)$ is the probability of the test result E_1 on the inverted disease state.

We consider the form of the equation (2) to be transparent because in this form we can pair it with its clinical interpretations [73,74]. A detailed clinical interpretation of the equation can be found in Appendix A. When the second diagnostic step E_2 is performed, the probability of the presence of the disease in the patient changes to:

$$P_2(D|E_2,E_1) = \frac{P(E_2,E_1|D)P_0(D)}{P(E_2,E_1|D)P_0(D) + P(E_2,E_1|\neg D)P_0(\neg D)}\,, \tag{3}$$

where the pair (E_2, E_1) represents the state of the *combined test*. The term $P_2(E_2,E_1|D)$ is the probability of the result of the combined test (E_2, E_1) conditional on whether the disease D is present or not. $P_2(E_2,E_1|\neg D)$ is an analogous probability, but under the condition that the inverted disease state is taken into account. Aspects of combining two diagnostic tests are described in detail, e.g. in [75].

A tempting approach is to simplify the equation (3) by assuming that the combined test (E_2, E_1) is the set of two independent tests E_2 and E_1. The assumption of independence often referred to as *naive* would transform equation (3) into the form:

$$P_2(D|E_2,E_1) = \frac{P(E_2|D)P(E_1|D)P_0(D)}{P(E_2|D)P(E_1|D)P_0(D) + P(E_2|\neg D)P(E_1|\neg D)P_0(\neg D)}\,,$$

where $P(E_2|D)$ and $P(E_2|\neg D)$ are the probabilities of the second test result E_2 depending on the disease state D and $\neg D$, respectively. If test independence is assumed, then it is possible to formally calculate the positive predictive value of these combined two tests with knowledge of the individual sensitivities and specificities of the tests and the prevalence of the disease. Unfortunately, this simplifying assumption, if not well substantiated, will lead to misleading results in a large number of cases, and the results obtained should not be considered valid.

4.7.2. Basic Prognostic Model

As an example of mathematical prognostic prediction, we present the Cox proportional hazards model. Cox model is widespread; it has become a sub-model in prognostic joint models [76] and a second stage in so-called two-stage models such as landmarking [12]. It has many variants and extensions [49]. It is challenged only by a model called discrete time logistic regression [66] developed in [51]. The original logistic regression compares the Cox model only when the cohort decline is not significant and the hazard rate function

can be approximated by a single value. It should be noted that if cohort attrition is the only concern, a two-step "smoothed" Cox method[4] can be used, such as in [77].

The following is not intended to compete with the explanations of the Cox model available in the current literature [48–50], but merely to provide a tangible example of a hazard rate function for the interested reader. The Cox proportional hazards model expresses *individualized hazard rate functions* from the statistics of the entire clinical trial cohort. The basic model input is that there are n patients indexed $i = 1, 2, \ldots n$ and each patient has p clinical parameters. The parameters of the ith patient can then be denoted as $x_{ji}, j = 1, 2, \ldots p$. These are the values that are recorded when a patient enters a clinical trial. The expression for the individualized hazard rate function $h_i(t)$ for ith patient has the form [48]:

$$h_i(t) = h_0(t) \exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_p x_{pi}) \,, \tag{4}$$

where $\beta_1, \beta_2, \ldots \beta_p$ are constants determined by the modeling process. Patient characteristics x_{ji} may be the results of his clinical tests or may represent his demographic data. The function $h_0(t)$ is the *basal hazard rate function*. The adjective *basal* means that it provides reference values for all individualized hazard rate functions. In the Cox proportional hazards model, the ratio between the values of the patient's hazard rate function and the basal hazard rate function does not change over time. The ratio is fully determined by the model constants $\boldsymbol{\beta}$ and values of patient characteristics x_{ji}. The sum of the values of the predictor x_{ji} multiplied by the coefficients β_j seen in the equation (4) is called the *linear term*. Determining the values of the $\boldsymbol{\beta}$ coefficients is a key modeling issue and the topic is discussed in Appendix B.

Patients' overall risk is often expressed as a patient's *risk score*. There are different ways of expressing its value, the SHFM risk score given in [77] seems most appropriate for the subsection describing the Cox model. The patient's risk score is expressed there in a very convenient way; the risk score called SHFM is simply the linear part of the equation (4). Other ways of defining risk scores can be found in the literature.

5. Overview of Heart Failure Prediction Models

The assumption of all the following prediction models is that the diagnosis of CHFS has already been made in the patient. We would like to repeat that all the publications reviewed here do not deal with *de-novo acute heart failure* [8,10,78]. All reviewed models assume already diagnosed chronic heart failure syndrome (CHFS) and predict acute decompensation (ADHF) or death. Prediction of *new-onset heart failure* or *incident heart failure* is also a separate topic and these models are reviewed in [79].

5.1. Telemedical Diagnostic Prediction of Acute Heart Failure Decompensation

We consider that the primary purpose of these works is early detection or diagnosis of acute decompensation (ADHF). Prediction models can be distinguished according to the criterion of whether an invasive or non-invasive method was used in relation to the patient. The prediction *object* is the early stage of ADHF. The earliest stage of detection for diagnosis is achieved by measuring the increase in pulmonary arterial pressure [16]. We consider, as already mentioned, that the pathological process of decompensation can occur several weeks before it is clearly manifested by the admission of the patient to the hospital.

5.1.1. Published Samples of Daily Monitored Telemedical Data

Weight change due to fluid retention is considered the most important predictor variable in the CHFS telemedicine monitoring system. Considerable work has been devoted to this matter of fact. To get a sense of the weight change of a patient before and after hospitalization with ADHF, one should review the real data or their averaged profiles, which can be

[4] If we accept the assumption of constancy of the hazard rate function in the Cox method, the survival curve has the form of a simple smooth exponential and does not contain any rips.

found in [54,62,80,81]. Daily intrathoracic bioimpedance data in the post-discharge period are presented in [82]. The daily dependence of intrathoracic impedance before and during hospitalization is presented in [83]. Implanted devices have also been used to collect daily data from CHFS and other types of cardiac patients. Their characteristics and averages are listed in [55,56].

5.1.2. Non-invasive Prediction Methods

Among the first attempts to create a predictive model of the clinical deterioration of a CHFS patient is work [84]. Data were collected using a patient weight record book. Zhang et al. [54] used a classification method originating from the financial industry called MACD. In relation to the input data structuring introduced in subsection *Temporal Characteristics of Input Data during Continuous Detection and Diagnosis*, we should say that their forward target window size was chosen to be 14 days; the optimized retrospective window size for the predictors was found to be 80 days. The method used, at least in our opinion, is capable of good prediction of the upcoming stage of CHFS deterioration, despite the authors' skepticism about the method. Their work influenced later works.

There is a brief review of non-invasive ADHF detection models included in [64] and will not be repeated here. To enlarge their list of models, we present three more. The first one is AHDF prediction using wavelet transform [53]. Their predictive continuous detection method is based on a *sliding window* approach and pattern identification. The development of four different predictor variables was used - and four different pairs of sensitivities and specificities were obtained. These four parameters were daily collected body weight, blood pressure, heart rate and respiration rate. The second work we would like to add to the list is [85], where a number of proposed features based on a single time-varying variable were tested as a basis for physiological signal detection and ADHF diagnosis. To achieve better alert signaling performance these signals were merged using a naive Bayesian method. The last non-invasive ADHF prediction method we will mention is the work of [86]. The performance of Bayesian online change point detection (BOCPD) and retrospective change point detection (RCPD) methods was evaluated. The former was better for events with a rapid onset, the latter for events that have slower gradual changes.

5.1.3. Prediction Models Using Implants

Companies Medtronix and Optivol are known for integrating patient monitoring devices into implantable cardioverter-defibrillators and similar devices. The patient's clinical parameters were monitored by various sensors. The sensor signals were combined by a Bayesian believe network [87] to obtain an decompensation prediction. Their method was later named TriageHF™ risk score [88]. A concurrent effort in predicting ADHF decompensations was made by Boston Scientific [89]. Patients had to have an implanted defibrillator for cardiac resynchronization therapy. Their HeartLogic™ multisensor index and alert algorithm provides a sensitive and timely predictor of impending ADHF decompensation. Details of their signal evaluation can be found in [90,91]. Both implant-based prediction technologies have been comprehensively analyzed and evaluated in [68].

5.1.4. Confirmatory Diagnosis of ADHF

The above methods provide us with only a preliminary diagnostic indication of decompensation. This prediction is followed by a detailed examination in a medical facility. Even then, the diagnosis of ADHF is not completely certain. The issue is addressed by a systematic review of sensitivities and specificities of various diagnostic parameters in [19].

5.2. Prognostic Prediction with Cross-Sectional Predictors

Prognostic tools of this type in the treatment of CHFS are recommended in [6, Chapter 4.8]. They should be used both for the prognosis of death and for the prognosis of hospitalization, but the effectiveness varies. A brief overview of the tools is also provided.

A survey of statistical models was carried out in 2008 [92]. During the study period (1988-2007), multivariate logistic regression and Cox proportional hazards regression were mainly used. Less than 15% of the publications use the χ^2 test only. A analogous survey was repeated in 2022 [70]. Similarly, Cox regression, logistic regression, and score methods were considered typical statistical models. In addition to statistical models, the authors also investigated machine learning models and we will mention them later. The review [93] also includes statistical models and machine learning models, the review done in [94] can also be noted.

Prognostic information about the risk of decompensation or death is often encapsulated in a simple scoring system. The developed scores vary in performance and have been compared in many publications. A comparison of the popular SHFM risk score and the MAGGIC score could be found in [95]. The SHFM score [77] is the linear part of the "smoothed" Cox proportional hazards model, where the model has been adjusted to use a constant hazard rate function. The MAGGIC [96] score is a converted Poisson regression model predictor. For further risk score performance comparisons, see e.g. in [97,98].

5.3. Advanced Statistical Modeling with Time-dependent Predictors

The modeling situation becomes unexpectedly complicated when the time parameter appears not only as an event parameter but also as a part of the predictor variable. In that case, two different timelines appear. The first comes from the time of events and the associated data is called *time-to-event data*. The second timeline is from when the predictor characteristics (e.g. biomarkers) were collected and the associated data are usually called *longitudinal data*. There are three basic approaches to dealing with this situation from a statistical modeling point of view [99,100]:

(i) a naive approach - simply use the obtained longitudinal data as predictors in models such as the Cox proportional hazards model,
(ii) two-stage modeling approach where longitudinal predictors are addressed first and time-to-event data are incorporated later. The most used model of this class seems to be the *landmarking model* [12], a generalized landmark model was recently introduced in [101],
(iii) true *joint model* approach, which consists of two models coupled by sharing random effects [76,99,102].

There is a literature that has compared the advantages of landmarking approach and joint modeling [100]. Comparison by simulation is done in [103,104]. These models are used for prognostic predictions, but so far we found only few articles dealing with chronic heart failure syndrome [105,106].

5.4. Other Advanced Statistical Models

The following are a group of prediction models one to be aware of, but which do not belong in any of the previous subsections. The first type of model includes the phenomenon of long-term changes in the entire population and in the health care system. This phenomenon results in a temporal and spatial shift of the model constants. Models need to be recalibrated and the effect of the change is called *calibration drift*. An overview of these models is given in [107]. [108] approaches the topic of model calibration in general.

Incorporating time-varying coefficients into the Cox model is considered an extension of it in [49,109,110]. An alternative to obtaining the hazard function by a model like the Cox model, was demonstrated in [51]. The model is presented in the review of early warning systems [66] and is labeled there as *discrete time logistic regression*.

5.5. A Machine Learning Approach to the Prognosis and Diagnosis in CHFS

Machine learning techniques (often referred to as artificial intelligence) also enter the field of clinical prediction modeling of CHFS. The application of machine learning methods to CHFS syndrome is freshly reviewed in [111–113]. It has become an excellent practice to compare the efficiency of machine learning classifiers with the efficiency of

established and well-researched logistic regression. We should note that recently the performance advantages of machine learning methods over traditional methods have been questioned [71]. In the following, we will divide our short review of machine learning in CHFS area into two parts.

5.5.1. Machine Learning for ADHF Detection and Diagnosis

Short-term prediction of hospitalization using a similarity-based machine learning (SBM) method was performed in [28]. Patients used a single wearable device during the clinical trial. The used positive window was 10 days long and corresponds to our forward target window. As a side note, the authors claim in the abstract that they have developed a prognostic algorithm to detect CHFS exacerbation. In this study, we would consider their prediction as part of the diagnostic process of decompensation, and we would prefer to call the algorithm a diagnostic detection algorithm.

The performance of seven machine learning methods was compared with the performance of logistic regression in [114]. The retrospective predictor window was assumed to be seven days, the forward target window was also chosen to be seven days. The Boruta method was used to eliminate insignificant predictors. The authors concluded that, among other methods, extreme gradient boost (XGBoost) method performs in the best way.

The performance of long short-term memory network (LSTM) was compared with logistic regression and the multi-layer percepron (MLP) method in [59]; LSTM was the best, logistic regression ranked second. The forward target window was chosen to be 30 days. Three groups of time-dependent predictor data were used. These groups were designated *fixed time* for demographic data, *diagnostic or episodic* for biomarker and medical examination data, and *high resolution* for vital signs data monitored on a frequent basis. Their method can be analogous to the previously discussed data fusion method.

5.5.2. Machine Learning for CHFS Prognosis

Two hundred and two statistical models were compared with 78 machine learning models in [70]. Random forests, support vector machine boosting, decision tree, MLP, and deep learning were listed among the machine learning methods. The authors concluded that ML models do not achieve significant benefits in event prediction. On the other hand, the authors of another comparison [115] concluded that machine learning classifiers perform better, but noted that ML prediction models should, as a rule, be reviewed using clinical modelling quality standards.

With a cohort of 30,687 adults, the performance of MLP, random forest, and XGBoost machine learning algorithms was compared with logistic regression in [116]. AUROC values were compared for 30-day, 90-day, and 365-day predictions for four different predictor engineering approaches. Except for the 90-day readmission, the XGBoost predictive models performed better than the other models. Prediction of CHFS readmission using LSTM for 30 days was reported in [117].

To conclude this subsection, we would like to add that the aforementioned work [70] also contains a critical comprehensive appreciation of machine learning efforts in CHFS modeling. Using the PROBAST tool, the survey authors concluded that currently, machine learning models generally have poorer clinical feasibility and reliability compared to statistical models.

6. Future Directions

The field of heart failure prediction research contains a number of publications that differ in the nature of input data, data processing, and prediction goals. The most elaborated is the prognostic prediction of the patient's death based on cross-sectional input data and it is a procedure ready for clinical deployment.

The area of continuous diagnosis, where the patient is monitored continuously over time, seems to be the least theoretically and practically explored. Digitization and lowering the price of medical measuring devices together with the development of telecommunica-

tion technologies enables obtaining medical information in the patient's home environment. From this point of view, these continuos diagnosis methods deserve attention. By structuring, grouping and reclassifying the input data in the section *Common Characteristics of Prediction Models* we made an attempt to indicate the direction of research in this area. Creating a classification scheme that provides a unifying view indicates possible improvements. We believe that the classification will be helpful in determining the theoretical and practical information maximum seen from the perspective of the therapist's analytical and decision-making processes.

Another finding from our review is that the vast majority of prediction publications, at least in the field of CHFS, focus on the prediction of adverse events such as death or hospital admission. In our opinion, the priority should go to the decision-making process of the therapist. In the field of telemedical CHFS, this could be optimization of decision-making when administering diuretics or outpatient up-titration.

We can reiterate here that there are significant differences between the US and European heart failure guidelines. We believe that the unified modeling approaches presented in the first two parts of this study could be helpful in solving this issue.

At the end of this section, we would like to note, that it seems that machine learning is accepting challenge from the community of medicine statisticians and starts to accept strictness and prudence of their modeling guidelines. Works are emerging that directly compare machine learning techniques with well-established methods such as logistic regression. There are also publications where machine learning is explicitly compared to statistical models. Once these challenges are met, machine learning may become an established clinical prediction technology.

7. Summary and Conclusion

The article provides introductory information on CHFS and its telemedicine, an overview of the basic common characteristics of predictive models, and finally a structured review of modeling literature primarily related to chronic heart failure syndrome. We summarized the information on CHFS found in the literature into a simplified discrete-state disease model condensed into two diagrams. In addition to the disease model we provide general desciption of inputs, outputs, and other characteristics of quantitative models. For orientation in the field of statistical modeling and clinical prediction, we have listed introductory literature and generally accepted guidelines.

We conducted a cross-sectional search of articles in the MEDLINE database of medical publications. During this process, we tried not to limit ourselves to a specific research group or direction of predictive research. Consequently, current prediction research is structured into four distinct research groups, each with slightly different methods and terminology. The first group of research focuses primarily on the engineering aspects of ADHF detection and diagnostics. The second group appears to be made up of medical statisticians, using well-established mainly prognostic prediction methods to maximize benefit to the medical audience. The third group uses advanced statistical methods to develop a patient prognosis with maximum use of time-dependent medical parameters. The last group, the machine learning group, tends to apply machine learning methods to the detection and diagnosis as well as prognosis of ADHF in a similar way to those used by the groups mentioned above.

It was only later that we realized that focusing on a specific area of CHFS has its limits in addition to its advantages. It is true that the issue of continuous remote monitoring present in the care of a patient with CHFS gave us a unique insight into the perspective of prediction methods. On the other hand, some directions seem to be marginal in the field of CHFS, and pointing out this fact has become one of the contents of this work.

8. Limitations of this Study

Given the primary purpose of providing a brief insight into the current state of clinical predictive modeling, we are aware of several limitations of this work. In the article, we did not explicitly deal with the process of training, validation and calibration of the prediction

model. We also did not address the connection of the prediction model to analytical therapeutic processes, decision-making therapeutic processes and the resulting value of information (VOI).

Funding: This publication was supported by the Operational Programme Integrated Infrastructure (OPII) for the project 313011BWH2: "InoCHF – Research and development in the field of innovative technologies in the management of patients with CHF", co-financed by the European Regional Development Fund.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The author declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

EHR	Electronic Health Record
UML	Unified Modeling Language
HF	Heart Failure
CHFS	Chronic Heart Failure Syndrome
ADHF	Acute Decompensation of Heart Failure
SHFM	Seattle Heart Failure Model
ER	Emergency Room
ICU	Intensive Care Unit
AUROC	Area Under the Receiver Operating Characteristics
ML	Machine Learning
MLP	Multi-Layer Perceptron
LSTM	Long Short-Term Memory
XGBoost	eXtreme Gradient Boosting

Appendix A. Explanatory material for the Bayes' equation

The clinical interpretation of the Bayes' equation (2) has four different forms depending on the state of the variables D and E_1. The left side of the equation can take the form that in clinical practice is called *positive predictive value* [73,74]. We can construct an explanatory table for all combinations of disease state D and test result E_1 in form of Table A1.

Table A1. Explanatory descriptions of Bayes' equation terms. $D = 1$ means the presence of the disease, $D = 0$ means the absence of the disease. $E_1 = 1$ means a positive test and $E_1 = 0$ means a negative test.

	$D = 1, E_1 = 1$	$D = 0, E_1 = 1$	$D = 1, E_1 = 0$	$D = 0, E_1 = 0$
$P_1(D\|E_1)$	Positive predictive value	False omission rate*	False discovery rate*	Negative predictive value
$P_0(D)$	Prevalence	1 - Prevalence	Prevalence	1 - Prevalence
$P_0(\neg D)$	1 - Prevalence	Prevalence	1 - Prevalence	Prevalence
$P(E_1\|D)$	Sensitivity	1 - Specificity	1 - Sensitivity	Specificity
$P(E_1\|\neg D)$	1 - Sensitivity	Sensitivity	Specificity	1 - Specificity
$P(E_1\|D)P_0(D)$	Probability of true positive	Probability of false positive	Probability of false negative	Probability of true negative
$P(E_1\|\neg D)P_0(\neg D)$	Probability of false positive	Probability of true positive	Probability of true negative	Probability of false negative

*It seems, that these terms has not been estabilished in clinical practice yet.

Supplementary information to the Table A1 is given in the Table A2. Definitions of established medical and statistical terms in the table were taken from the basic literature.

Table A2. Quantitative definition of key diagnostic concepts.

Concept name	Brief description
Prevalence	Proportion of a defined group in the population having a disease at one point in time
Sensitivity	Rate of positive responses in a test from persons with a specific disease, true positive rate
Specificity	Rate of negative responses in a test from persons free from a disease, true negative rate
True positives	Number of cases in population correctly identified as diseased
False positives	Number of cases in population incorrectly identified as diseased, type I error
True negatives	Number of cases in population correctly identified as healthy
False negatives	Number of cases in population incorrectly identified as healthy, type II error

Appendix B. Determining Coefficients in Cox Regression Using the Maximum Likelihood Estimation Method.

Cox's *partial likelihood function* $L(\boldsymbol{\beta})$ is used to determine the values of the $\boldsymbol{\beta}$ coefficients in the equation (4). The patient leaves the clinical trial at time t_i either because of the occurrence of the investigated event or for another reason usually included under the term censoring. The coefficients $\boldsymbol{\beta}$ are determined by maximizing the value $L(\boldsymbol{\beta})$ whose logarithm is given by [48]

$$\log L(\boldsymbol{\beta}) = \sum_{i=1}^{n} \delta_i \left\{ \sum_{j=1}^{p} \beta_j x_{ji}(t_i) - \log \sum_{l \in R(t_i)} \exp\left(\sum_{j=1}^{p} \beta_j x_{jl}(t_i) \right) \right\},$$

where δ_i takes on values of zero for censored patients and unity for a patient experiencing the investigated event. We could see that if the *ith* patient is censored, δ_i nullifies its contribution in overall summation. Censored patient data values are only *partially* used in the internal sum and are accounted for through index l. The inner sum applies to all patients in subset $R(t_i)$, which is the subset composed of patients who did not experience an event and were uncensored just before time t_i [48]. See [48–50] for more details.

References

1. Celi, L.A.; Charlton, P.; Ghassemi, M.M.; Johnson, A.; Komorowski, M.; Marshall, D.; Naumann, T.; Paik, K.; Pollard, T.J.; Raffa, J.; et al. *Secondary analysis of electronic health records*; MIT Critical Data, Springer Nature, 2016.
2. Zhou, Z.R.; Wang, W.W.; Li, Y.; Jin, K.R.; Wang, X.Y.; Wang, Z.W.; Chen, Y.S.; Wang, S.J.; Hu, J.; Zhang, H.N.; et al. In-depth mining of clinical data: the construction of clinical prediction model with R. *Annals of translational medicine* **2019**, *7*.
3. Smith, H.; Sweeting, M.; Morris, T.; Crowther, M.J. A scoping methodological review of simulation studies comparing statistical and machine learning approaches to risk prediction for time-to-event data. *Diagnostic and Prognostic Research* **2022**, *6*, 1–15.
4. Steyerberg, E.W.; Vergouwe, Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. *European heart journal* **2014**, *35*, 1925–1931.
5. Steyerberg, E.W. *Clinical prediction models: A Practical Approach to Development, Validation, and Updating*, 2nd ed.; Springer, 2009.
6. Heidenreich, P.A.; Bozkurt, B.; Aguilar, D.; Allen, L.A.; Byun, J.J.; Colvin, M.M.; Deswal, A.; Drazner, M.H.; Dunlay, S.M.; Evers, L.R.; et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Journal of the American College of Cardiology* **2022**, *79*, e263–e421.
7. McDonagh, T.A.; Metra, M.; Adamo, M.; Gardner, R.S.; Baumbach, A.; Böhm, M.; Burri, H.; Butler, J.; Čelutkienė, J.; Chioncel, O.; et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: Developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. *European heart journal* **2021**, *42*, 3599–3726.
8. Kurmani, S.; Squire, I. Acute heart failure: definition, classification and epidemiology. *Current heart failure reports* **2017**, *14*, 385–392.

9. Straw, S.; Napp, A.; Witte, K.K. 'Acute Heart Failure': Should We Abandon the Term Altogether? *Current Heart Failure Reports* **2022**, *19*, 425–434.
10. Hummel, A.; Empen, K.; Dörr, M.; Felix, S.B. De novo acute heart failure and acutely decompensated chronic heart failure. *Deutsches Ärzteblatt International* **2015**, *112*, 298.
11. Mann, D.L.; Bristow, M.R. Mechanisms and models in heart failure: the biomechanical model and beyond. *Circulation* **2005**, *111*, 2837–2849.
12. Houwelingen, H.; Putter, H. *Dynamic prediction in clinical survival analysis*; CRC Press, 2011.
13. Januzzi Jr, J.L.; Butler, J. The importance of worsening heart failure: hiding in plain sight. *Journal of the American College of Cardiology* **2022**, *80*, 123–125.
14. Schiff, G.D.; Fung, S.; Speroff, T.; McNutt, R.A. Decompensated heart failure: symptoms, patterns of onset, and contributing factors. *The American journal of medicine* **2003**, *114*, 625–630.
15. Ho, J.E.; Lyass, A.; Lee, D.S.; Vasan, R.S.; Kannel, W.B.; Larson, M.G.; Levy, D. Predictors of new-onset heart failure: differences in preserved versus reduced ejection fraction. *Circulation: heart failure* **2013**, *6*, 279–286.
16. Adamson, P.B. Pathophysiology of the transition from chronic compensated and acute decompensated heart failure: new insights from continuous monitoring devices. *Current heart failure reports* **2009**, *6*, 287–292.
17. Dworzynski, K.; Roberts, E.; Ludman, A.; Mant, J. Diagnosing and managing acute heart failure in adults: summary of NICE guidance. *BmJ* **2014**, *349*.
18. Ramani, G.V.; Uber, P.A.; Mehra, M.R. Chronic heart failure: contemporary diagnosis and management. In Proceedings of the Mayo Clinic Proceedings. Elsevier, 2010, Vol. 85, pp. 180–195.
19. Martindale, J.L.; Wakai, A.; Collins, S.P.; Levy, P.D.; Diercks, D.; Hiestand, B.C.; Fermann, G.J.; Desouza, I.; Sinert, R. Diagnosing acute heart failure in the emergency department: a systematic review and meta-analysis. *Academic emergency medicine* **2016**, *23*, 223–242.
20. Kellett, J.; Sebat, F. Make vital signs great again–A call for action. *European journal of internal medicine* **2017**, *45*, 13–19.
21. Drews, T.E.; Laukkanen, J.; Nieminen, T. Non-invasive home telemonitoring in patients with decompensated heart failure: a systematic review and meta-analysis. *ESC Heart Failure* **2021**, *8*, 3696–3708.
22. Kropf, M.; Modre-Osprian, R.; Hayn, D.; Fruhwald, F.; Schreier, G. Telemonitoring in heart failure patients with clinical decision support to optimize medication doses based on guidelines. In Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2014, pp. 3168–3171.
23. Louis, A.A.; Turner, T.; Gretton, M.; Baksh, A.; Cleland, J.G. A systematic review of telemonitoring for the management of heart failure. *European journal of heart failure* **2003**, *5*, 583–590.
24. Zito, A.; Princi, G.; Romiti, G.F.; Galli, M.; Basili, S.; Liuzzo, G.; Sanna, T.; Restivo, A.; Ciliberti, G.; Trani, C.; et al. Device-based remote monitoring strategies for congestion-guided management of patients with heart failure: a systematic review and meta-analysis. *European Journal of Heart Failure* **2022**, *24*, 2333–2341.
25. Curtain, J.P.; Lee, M.M.; McMurray, J.J.; Gardner, R.S.; Petrie, M.C.; Jhund, P.S. Efficacy of implantable haemodynamic monitoring in heart failure across ranges of ejection fraction: a systematic review and meta-analysis. *Heart* **2023**, *109*, 823–831.
26. McGee, M.J.; Ray, M.; Brienesse, S.C.; Sritharan, S.; Boyle, A.J.; Jackson, N.; Leitch, J.W.; Sverdlov, A.L. Remote monitoring in patients with heart failure with cardiac implantable electronic devices: a systematic review and meta-analysis. *Open Heart* **2022**, *9*, e002096.
27. Straw, S.; Witte, K.K. Remote monitoring in heart failure: it's the data you collect and what you do with them, 2023.
28. Stehlik, J.; Schmalfuss, C.; Bozkurt, B.; Nativi-Nicolau, J.; Wohlfahrt, P.; Wegerich, S.; Rose, K.; Ray, R.; Schofield, R.; Deswal, A.; et al. Continuous wearable monitoring analytics predict heart failure hospitalization: the LINK-HF multicenter study. *Circulation: Heart Failure* **2020**, *13*, e006513.
29. Lin, M.h.; Yuan, W.l.; Huang, T.c.; Zhang, H.f.; Mai, J.t.; Wang, J.f. Clinical effectiveness of telemedicine for chronic heart failure: a systematic review and meta-analysis. *Journal of Investigative Medicine* **2017**, *65*, 899–911.
30. Kuan, P.X.; Chan, W.K.; Ying, D.K.F.; Rahman, M.A.A.; Peariasamy, K.M.; Lai, N.M.; Mills, N.L.; Anand, A. Efficacy of telemedicine for the management of cardiovascular disease: a systematic review and meta-analysis. *The Lancet Digital Health* **2022**, *4*, e676–e691.
31. Alvarez, P.; Sianis, A.; Brown, J.; Ali, A.; Briasoulis, A. Chronic disease management in heart failure: focus on telemedicine and remote monitoring. *Reviews in cardiovascular medicine* **2021**, *22*, 403–413.
32. Umeh, C.A.; Torbela, A.; Saigal, S.; Kaur, H.; Kazourra, S.; Gupta, R.; Shah, S. Telemonitoring in heart failure patients: Systematic review and meta-analysis of randomized controlled trials. *World Journal of Cardiology* **2022**, *14*, 640.
33. Rebolledo Del Toro, M.; Herrera Leano, N.M.; Barahona-Correa, J.E.; Munoz Velandia, O.M.; Fernández Ávila, D.G.; García Peña, Á.A. Effectiveness of mobile telemonitoring applications in heart failure patients: systematic review of literature and meta-analysis. *Heart Failure Reviews* **2023**, *28*, 431–452.
34. Koehler, F.; Koehler, K.; Deckwart, O.; Prescher, S.; Wegscheider, K.; Kirwan, B.A.; Winkler, S.; Vettorazzi, E.; Bruch, L.; Oeff, M.; et al. Efficacy of telemedical interventional management in patients with heart failure (TIM-HF2): a randomised, controlled, parallel-group, unmasked trial. *The Lancet* **2018**, *392*, 1047–1057.
35. Chaudhry, S.I.; Mattera, J.A.; Curtis, J.P.; Spertus, J.A.; Herrin, J.; Lin, Z.; Phillips, C.O.; Hodshon, B.V.; Cooper, L.S.; Krumholz, H.M. Telemonitoring in patients with heart failure. *New England Journal of Medicine* **2010**, *363*, 2301–2309.

36. Galinier, M.; Roubille, F.; Berdague, P.; Brierre, G.; Cantie, P.; Dary, P.; Ferradou, J.M.; Fondard, O.; Labarre, J.P.; Mansourati, J.; et al. Telemonitoring versus standard care in heart failure: a randomised multicentre trial. *European Journal of Heart Failure* **2020**, *22*, 985–994.
37. Ong, M.K.; Romano, P.S.; Edgington, S.; Aronow, H.U.; Auerbach, A.D.; Black, J.T.; De Marco, T.; Escarce, J.J.; Evangelista, L.S.; Hanna, B.; et al. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the better effectiveness after transition–heart failure (BEAT-HF) randomized clinical trial. *JAMA internal medicine* **2016**, *176*, 310–318.
38. Black, J.T.; Romano, P.S.; Sadeghi, B.; Auerbach, A.D.; Ganiats, T.G.; Greenfield, S.; Kaplan, S.H.; Ong, M.K.; Group, B.H.R. A remote monitoring and telephone nurse coaching intervention to reduce readmissions among patients with heart failure: study protocol for the Better Effectiveness After Transition-Heart Failure (BEAT-HF) randomized controlled trial. *Trials* **2014**, *15*, 1–11.
39. Binuya, M.; Engelhardt, E.; Schats, W.; Schmidt, M.; Steyerberg, E. Methodological guidance for the evaluation and updating of clinical prediction models: a systematic review. *BMC Medical Research Methodology* **2022**, *22*, 316.
40. Kappen, T.H.; van Klei, W.A.; van Wolfswinkel, L.; Kalkman, C.J.; Vergouwe, Y.; Moons, K.G. Evaluating the impact of prediction models: lessons learned, challenges, and recommendations. *Diagnostic and prognostic research* **2018**, *2*, 1–11.
41. Duran, A.; De Anda-Duran, I.; Ventura, H.O. The era of heart failure risk prediction models, is it time to test their utility? *International Journal of Cardiology* **2022**, *352*, 98–99.
42. Collins, G.S.; Reitsma, J.B.; Altman, D.G.; Moons, K.G. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) the TRIPOD statement. *Circulation* **2015**, *131*, 211–219.
43. Moons, K.G.; Altman, D.G.; Reitsma, J.B.; Ioannidis, J.P.; Macaskill, P.; Steyerberg, E.W.; Vickers, A.J.; Ransohoff, D.F.; Collins, G.S. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Annals of internal medicine* **2015**, *162*, W1–W73.
44. Wolff, R.F.; Moons, K.G.; Riley, R.D.; Whiting, P.F.; Westwood, M.; Collins, G.S.; Reitsma, J.B.; Kleijnen, J.; Mallett, S.; Group†, P. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. *Annals of internal medicine* **2019**, *170*, 51–58.
45. Moons, K.G.; Wolff, R.F.; Riley, R.D.; Whiting, P.F.; Westwood, M.; Collins, G.S.; Reitsma, J.B.; Kleijnen, J.; Mallett, S. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. *Annals of internal medicine* **2019**, *170*, W1–W33.
46. Leisman, D.E.; Harhay, M.O.; Lederer, D.J.; Abramson, M.; Adjei, A.A.; Bakker, J.; Ballas, Z.K.; Barreiro, E.; Bell, S.C.; Bellomo, R.; et al. Development and reporting of prediction models: guidance for authors from editors of respiratory, sleep, and critical care journals. *Critical care medicine* **2020**, *48*, 623.
47. Ruppert, M.M.; Loftus, T.J.; Small, C.; Li, H.; Ozrazgat-Baslanti, T.; Balch, J.; Holmes, R.; Tighe, P.J.; Upchurch Jr, G.R.; Efron, P.A.; et al. Predictive Modeling for Readmission to Intensive Care: A Systematic Review. *Critical Care Explorations* **2023**, *5*.
48. Collett, D. *Modelling Survival Data in Medical Research*, 3rd ed.; CRC Press, 2015.
49. Therneau, T.; Grambsch, P.M. *Modelling Survival Data: Extending the Cox Model*, 1st ed.; Springer-Verlag, 2000.
50. Liu, X. *Survival analysis: models and applications*; John Wiley & Sons, 2012.
51. Efron, B. Logistic regression, survival analysis, and the Kaplan-Meier curve. *Journal of the American statistical Association* **1988**, *83*, 414–425.
52. Ahmad, T.; Munir, A.; Bhatti, S.H.; Aftab, M.; Raza, M.A. Survival analysis of heart failure patients: A case study. *PloS one* **2017**, *12*, e0181001.
53. Henriques, J.; Carvalho, P.; Paredes, S.; Rocha, T.; Habetha, J.; Antunes, M.; Morais, J. Prediction of heart failure decompensation events by trend analysis of telemonitoring data. *IEEE Journal of Biomedical and Health Informatics* **2014**, *19*, 1757–1769.
54. Zhang, J.; Goode, K.M.; Cuddihy, P.E.; Cleland, J.G.; Investigators, T.H. Predicting hospitalization due to worsening heart failure using daily weight measurement: analysis of the Trans-European Network-Home-Care Management System (TEN-HMS) study. *European journal of heart failure* **2009**, *11*, 420–427.
55. Gardner, R.S.; Thakur, P.; Hammill, E.F.; Nair, D.G.; Eldadah, Z.; Stančák, B.; Ferrick, K.; Sriratanasathavorn, C.; Duray, G.Z.; Wariar, R.; et al. Multiparameter diagnostic sensor measurements during clinically stable periods and worsening heart failure in ambulatory patients. *ESC heart failure* **2021**, *8*, 1571–1581.
56. Zile, M.R.; Kahwash, R.; Sarkar, S.; Koehler, J.; Butler, J. Temporal characteristics of device-based individual and integrated risk metrics in patients with chronic heart failure. *Heart Failure* **2023**, *11*, 143–156.
57. Alvarez, C.A.; Clark, C.A.; Zhang, S.; Halm, E.A.; Shannon, J.J.; Girod, C.E.; Cooper, L.; Amarasingham, R. Predicting out of intensive care unit cardiopulmonary arrest or death using electronic medical record data. *BMC medical informatics and decision making* **2013**, *13*, 1–11.
58. Churpek, M.M.; Yuen, T.C.; Winslow, C.; Robicsek, A.A.; Meltzer, D.O.; Gibbons, R.D.; Edelson, D.P. Multicenter development and validation of a risk stratification tool for ward patients. *American journal of respiratory and critical care medicine* **2014**, *190*, 649–655.
59. Fahimi, F.; Guo, Y.; Tong, S.C.; Ng, A.; Bing, S.O.Y.; Choo, B.; Weiliang, H.; Lee, S.; Ramasamy, S.; Chow, W.L.; et al. A Vital Signs Telemonitoring Programme Improves the Dynamic Prediction of Readmission Risk in Patients with Heart Failure. In Proceedings of the AMIA Annual Symposium Proceedings. American Medical Informatics Association, 2020, Vol. 2020, p. 432.
60. Valko, M.; Hauskrecht, M. Feature importance analysis for patient management decisions. *Studies in health technology and informatics* **2010**, *160*, 861.

61. Sarkar, S.; Koehler, J. A dynamic risk score to identify increased risk for heart failure decompensation. *IEEE Transactions on Biomedical Engineering* **2012**, *60*, 147–150.
62. Joshi, R.; Gyllensten, I.C. Changes in daily measures of blood pressure and heart rate improve weight-based detection of heart failure deterioration in patients on telemonitoring. *IEEE Journal of Biomedical and Health Informatics* **2018**, *23*, 1041–1048.
63. Ledwidge, M.; Mcdonald, K. Evaluation of a Subject's Weight, 2012. US Patent 20120330683-A1.
64. Larburu, N.; Artetxe, A.; Escolar, V.; Lozano, A.; Kerexeta, J. Artificial intelligence to prevent mobile heart failure patients decompensation in real time: monitoring-based predictive model. *Mobile Information Systems* **2018**, *2018*, 1–11.
65. Van Smeden, M.; Reitsma, J.B.; Riley, R.D.; Collins, G.S.; Moons, K.G. Clinical prediction models: diagnosis versus prognosis. *Journal of clinical epidemiology* **2021**, *132*, 142–145.
66. Fu, L.H.; Schwartz, J.; Moy, A.; Knaplund, C.; Kang, M.J.; Schnock, K.O.; Garcia, J.P.; Jia, H.; Dykes, P.C.; Cato, K.; et al. Development and validation of early warning score system: A systematic literature review. *Journal of biomedical informatics* **2020**, *105*, 103410.
67. Churpek, M.M.; Yuen, T.C.; Park, S.Y.; Gibbons, R.; Edelson, D.P. Using electronic health record data to develop and validate a prediction model for adverse outcomes on the wards. *Critical care medicine* **2014**, *42*, 841.
68. Assa, S.; Vernooy, K.; van Stipdonk, A.M. Cardiovascular Implantable Electronic Devices Enabled Remote Heart Failure Monitoring; What We Have Learned and Where to Go Next. *Journal of Cardiovascular Development and Disease* **2023**, *10*, 152.
69. Breiman, L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical science* **2001**, *16*, 199–231.
70. Sun, Z.; Dong, W.; Shi, H.; Ma, H.; Cheng, L.; Huang, Z. Comparing machine learning models and statistical models for predicting heart failure events: a systematic review and meta-analysis. *Frontiers in Cardiovascular Medicine* **2022**, *9*, 812276.
71. Christodoulou, E.; Ma, J.; Collins, G.S.; Steyerberg, E.W.; Verbakel, J.Y.; Van Calster, B. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. *Journal of clinical epidemiology* **2019**, *110*, 12–22.
72. Austin, P.C.; Harrell Jr, F.E.; Steyerberg, E.W. Predictive performance of machine and statistical learning methods: impact of data-generating processes on external validity in the "large N, small p" setting. *Statistical methods in medical research* **2021**, *30*, 1465–1483.
73. Bours, M.J. Bayes' rule in diagnosis. *Journal of Clinical Epidemiology* **2021**, *131*, 158–160.
74. Altman, D.G.; Bland, J.M. Statistics Notes: Diagnostic tests 2: predictive values. *Bmj* **1994**, *309*, 102.
75. Marshall, R.J. The predictive value of simple rules for combining two diagnostic tests. *Biometrics* **1989**, pp. 1213–1222.
76. Papageorgiou, G.; Mauff, K.; Tomer, A.; Rizopoulos, D. An overview of joint modeling of time-to-event and longitudinal outcomes. *Annual review of statistics and its application* **2019**, *6*, 223–240.
77. Levy, W.C.; Mozaffarian, D.; Linker, D.T.; Sutradhar, S.C.; Anker, S.D.; Cropp, A.B.; Anand, I.; Maggioni, A.; Burton, P.; Sullivan, M.D.; et al. The Seattle Heart Failure Model: prediction of survival in heart failure. *Circulation* **2006**, *113*, 1424–1433.
78. Raffaello, W.M.; Henrina, J.; Huang, I.; Lim, M.A.; Suciadi, L.P.; Siswanto, B.B.; Pranata, R. Clinical characteristics of de novo heart failure and acute decompensated chronic heart failure: are they distinctive phenotypes that contribute to different outcomes? *Cardiac Failure Review* **2021**, *7*.
79. Sahle, B.W.; Owen, A.J.; Chin, K.L.; Reid, C.M. Risk prediction models for incident heart failure: a systematic review of methodology and model performance. *Journal of cardiac failure* **2017**, *23*, 680–687.
80. Chaudhry, S.I.; Wang, Y.; Concato, J.; Gill, T.M.; Krumholz, H.M. Patterns of weight change preceding hospitalization for heart failure. *Circulation* **2007**, *116*, 1549–1554.
81. Ledwidge, M.T.; O'Hanlon, R.; Lalor, L.; Travers, B.; Edwards, N.; Kelly, D.; Voon, V.; McDonald, K.M. Can individualized weight monitoring using the HeartPhone algorithm improve sensitivity for clinical deterioration of heart failure? *European journal of heart failure* **2013**, *15*, 447–455.
82. Darling, C.E.; Dovancescu, S.; Saczynski, J.S.; Riistama, J.; Kuniyoshi, F.S.; Rock, J.; Meyer, T.E.; McManus, D.D. Bioimpedance-based heart failure deterioration prediction using a prototype fluid accumulation vest-mobile phone dyad: an observational study. *JMIR cardio* **2017**, *1*, e6057.
83. Yu, C.M.; Wang, L.; Chau, E.; Chan, R.H.W.; Kong, S.L.; Tang, M.O.; Christensen, J.; Stadler, R.W.; Lau, C.P. Intrathoracic impedance monitoring in patients with heart failure: correlation with fluid status and feasibility of early warning preceding hospitalization. *Circulation* **2005**, *112*, 841–848.
84. Lewin, J.; Ledwidge, M.; O'Loughlin, C.; McNally, C.; McDonald, K. Clinical deterioration in established heart failure: what is the value of BNP and weight gain in aiding diagnosis? *European journal of heart failure* **2005**, *7*, 953–957.
85. Joshi, R.; Gyllensten, I.C. Changes in daily measures of blood pressure and heart rate improve weight-based detection of heart failure deterioration in patients on telemonitoring. *IEEE Journal of Biomedical and Health Informatics* **2018**, *23*, 1041–1048.
86. Javed, F.; Farrugia, S.; Colefax, M.; Schindhelm, K. Early warning of acute decompensation in heart failure patients using a noncontact measure of stability index. *IEEE Transactions on Biomedical Engineering* **2015**, *63*, 438–448.
87. Cowie, M.R.; Sarkar, S.; Koehler, J.; Whellan, D.J.; Crossley, G.H.; Tang, W.H.W.; Abraham, W.T.; Sharma, V.; Santini, M. Development and validation of an integrated diagnostic algorithm derived from parameters monitored in implantable devices for identifying patients at risk for heart failure hospitalization in an ambulatory setting. *European Heart Journal* **2013**, *34*, 2472–2480.
88. Zile, M.R.; Koehler, J.; Sarkar, S.; Butler, J. Prediction of worsening heart failure events and all-cause mortality using an individualized risk stratification strategy. *ESC Heart Failure* **2020**, *7*, 4277–4289.

89. Boehmer, J.P.; Hariharan, R.; Devecchi, F.G.; Smith, A.L.; Molon, G.; Capucci, A.; An, Q.; Averina, V.; Stolen, C.M.; Thakur, P.H.; et al. A multisensor algorithm predicts heart failure events in patients with implanted devices: results from the MultiSENSE study. *JACC: Heart Failure* **2017**, *5*, 216–225.
90. Boehmer, J.P.; Sriratanasathavorn, C.; Fisher, J.; Bransford, P.; Chan, R.; Sweeney, R.J.; Ahmed, R.; Zhang, Y.; Averina, V.; An, Q.; et al. Heart failure diagnostics sensor measurements change prior to heart failure decompensation events. *Journal of Cardiac Failure* **2017**, *23*, S65.
91. López-Azor, J.C.; de la Torre, N.; Carmena, M.D.G.C.; Pérez, P.C.; Munera, C.; MarcoClement, I.; León, R.C.; Álvarez-García, J.; Pachón, M.; Ynsaurriaga, F.A.; et al. Clinical Utility of HeartLogic, a Multiparametric Telemonitoring System, in Heart Failure. *Cardiac Failure Review* **2022**, *8*.
92. Ross, J.S.; Mulvey, G.K.; Stauffer, B.; Patlolla, V.; Bernheim, S.M.; Keenan, P.S.; Krumholz, H.M. Statistical models and patient predictors of readmission for heart failure: a systematic review. *Archives of internal medicine* **2008**, *168*, 1371–1386.
93. Artetxe, A.; Beristain, A.; Grana, M. Predictive models for hospital readmission risk: A systematic review of methods. *Computer methods and programs in biomedicine* **2018**, *164*, 49–64.
94. Liu, J.; Liu, P.; Lei, M.R.; Zhang, H.W.; You, A.L.; Luan, X.R. Readmission Risk Prediction Model for Patients with Chronic Heart Failure: A Systematic Review and Meta-Analysis. *Iranian Journal of Public Health* **2022**, *51*, 1481.
95. Rich, J.D.; Burns, J.; Freed, B.H.; Maurer, M.S.; Burkhoff, D.; Shah, S.J. Meta-Analysis Global Group in Chronic (MAGGIC) heart failure risk score: validation of a simple tool for the prediction of morbidity and mortality in heart failure with preserved ejection fraction. *Journal of the American Heart Association* **2018**, *7*, e009594.
96. Pocock, S.J.; Ariti, C.A.; McMurray, J.J.; Maggioni, A.; Køber, L.; Squire, I.B.; Swedberg, K.; Dobson, J.; Poppe, K.K.; Whalley, G.A.; et al. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. *European heart journal* **2013**, *34*, 1404–1413.
97. Canepa, M.; Fonseca, C.; Chioncel, O.; Laroche, C.; Crespo-Leiro, M.G.; Coats, A.J.; Mebazaa, A.; Piepoli, M.F.; Tavazzi, L.; Maggioni, A.P.; et al. Performance of prognostic risk scores in chronic heart failure patients enrolled in the European Society of Cardiology Heart Failure Long-Term Registry. *JACC: Heart Failure* **2018**, *6*, 452–462.
98. Codina, P.; Lupón, J.; Borrellas, A.; Spitaleri, G.; Cediel, G.; Domingo, M.; Simpson, J.; Levy, W.C.; Santiago-Vacas, E.; Zamora, E.; et al. Head-to-head comparison of contemporary heart failure risk scores. *European Journal of Heart Failure* **2021**, *23*, 2035–2044.
99. Gould, L.A.; Boye, M.E.; Crowther, M.J.; Ibrahim, J.G.; Quartey, G.; Micallef, S.; Bois, F.Y. Joint modeling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modeling working group. *Statistics in medicine* **2015**, *34*, 2181–2195.
100. Gao, F.; Luo, J.; Liu, J.; Wan, F.; Wang, G.; Gordon, M.; Xiong, C. Comparing statistical methods in assessing the prognostic effect of biomarker variability on time-to-event clinical outcomes. *BMC Medical Research Methodology* **2022**, *22*, 201.
101. Yao, Y.; Li, L.; Astor, B.; Yang, W.; Greene, T. Predicting the risk of a clinical event using longitudinal data: the generalized landmark analysis. *BMC Medical Research Methodology* **2023**, *23*, 5.
102. Rizopoulos, D. *Joint models for longitudinal and time-to-event data: With applications in R*; CRC press, 2012.
103. Suresh, K.; Taylor, J.M.; Spratt, D.E.; Daignault, S.; Tsodikov, A. Comparison of joint modeling and landmarking for dynamic prediction under an illness-death model. *Biometrical Journal* **2017**, *59*, 1277–1300.
104. Li, W.; Li, L.; Astor, B.C. A comparison of two approaches to dynamic prediction: Joint modeling and landmark modeling. *Statistics in Medicine* **2023**, *42*, 2101–2115.
105. van Vark, L.C.; Lesman-Leegte, I.; Baart, S.J.; Postmus, D.; Pinto, Y.M.; Orsel, J.G.; Westenbrink, B.D.; Brunner-la Rocca, H.P.; van Miltenburg, A.J.; Boersma, E.; et al. Prognostic value of serial ST2 measurements in patients with acute heart failure. *Journal of the American College of Cardiology* **2017**, *70*, 2378–2388.
106. Canepa, M.; Siri, G.; Puntoni, M.; Latini, R.; Tavazzi, L.; Maggioni, A.P. Testing longitudinal data for prognostication in ambulatory heart failure patients with reduced ejection fraction. A proof of principle from the GISSI-HF database. *International Journal of Cardiology* **2020**, *313*, 89–96.
107. Jenkins, D.A.; Sperrin, M.; Martin, G.P.; Peek, N. Dynamic models to predict health outcomes: current status and methodological challenges. *Diagnostic and prognostic research* **2018**, *2*, 1–9.
108. Van Calster, B.; McLernon, D.J.; Van Smeden, M.; Wynants, L.; Steyerberg, E.W. Calibration: the Achilles heel of predictive analytics. *BMC medicine* **2019**, *17*, 1–7.
109. Therneau, T.; Crowson, C.; Atkinson, E. Using time dependent covariates and time dependent coefficients in the cox model. *Survival Vignettes* **2017**, *2*, 1–25.
110. Zhang, Z.; Reinikainen, J.; Adeleke, K.A.; Pieterse, M.E.; Groothuis-Oudshoorn, C.G. Time-varying covariates and coefficients in Cox regression models. *Annals of translational medicine* **2018**, *6*.
111. Croon, P.; Selder, J.; Allaart, C.; Bleijendaal, H.; Chamuleau, S.; Hofstra, L.; Išgum, I.; Ziesemer, K.; Winter, M. Current state of artificial intelligence-based algorithms for hospital admission prediction in patients with heart failure: a scoping review. *European Heart Journal-Digital Health* **2022**, *3*, 415–425.
112. Gautam, N.; Ghanta, S.N.; Mueller, J.; Mansour, M.; Chen, Z.; Puente, C.; Ha, Y.M.; Tarun, T.; Dhar, G.; Sivakumar, K.; et al. Artificial Intelligence, Wearables and Remote Monitoring for Heart Failure: Current and Future Applications. *Diagnostics* **2022**, *12*, 2964.

113. Mortazavi, B.J.; Downing, N.S.; Bucholz, E.M.; Dharmarajan, K.; Manhapra, A.; Li, S.X.; Negahban, S.N.; Krumholz, H.M. Analysis of machine learning techniques for heart failure readmissions. *Circulation: Cardiovascular Quality and Outcomes* **2016**, *9*, 629–640.
114. Kerexeta, J.; Larburu, N.; Escolar, V.; Lozano-Bahamonde, A.; Macía, I.; Beristain Iraola, A.; Graña, M. Prediction and analysis of heart failure decompensation events based on telemonitored data and artificial intelligence methods. *Journal of Cardiovascular Development and Disease* **2023**, *10*, 48.
115. Shin, S.; Austin, P.C.; Ross, H.J.; Abdel-Qadir, H.; Freitas, C.; Tomlinson, G.; Chicco, D.; Mahendiran, M.; Lawler, P.R.; Billia, F.; et al. Machine learning vs. conventional statistical models for predicting heart failure readmission and mortality. *ESC heart failure* **2021**, *8*, 106–115.
116. Ru, B.; Tan, X.; Liu, Y.; Kannapur, K.; Ramanan, D.; Kessler, G.; Lautsch, D.; Fonarow, G. Comparison of machine learning algorithms for predicting hospital readmissions and worsening heart failure events in patients with heart failure with reduced ejection fraction: Modeling study. *JMIR Formative Research* **2023**, *7*, e41775.
117. Ashfaq, A.; Sant'Anna, A.; Lingman, M.; Nowaczyk, S. Readmission prediction using deep learning on electronic health records. *Journal of biomedical informatics* **2019**, *97*, 103256.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.