Shap. Mem. Superelasticity
https://doi.org/10.1007/s40830-023-00421-5

REVIEW

Advancing Nitinol Implant Design and Simulation Through Data-Driven Methodologies

Harshad M. Paranjape[1,2]

Received: 3 August 2022 / Revised: 15 December 2022 / Accepted: 18 January 2023

Abstract Recent advances in the Data Science methods for acquiring and analyzing large amounts of materials deformation data have the potential to tremendously benefit Nitinol (Nickel–Titanium shape memory alloy) implant design and simulation. We review some of these data-driven methodologies and provide a perspective on adapting these techniques to Nitinol design and simulation. We organize the review in a three-tiered approach. The methods in the first tier relate to data acquisition. We review methods for acquiring full-field deformation data from implants and methods for quantifying uncertainty in such data. The second-tier methods relate to combining data from multiple sources to gain a holistic understanding of complex deformation phenomena such as fatigue. Methods in the third tier relate to making data-driven simulation of the deformation response of Nitinol. A wide adaption of these methods by the Nitinol cardiovascular implant community may be facilitated by building consensus on best practices and open exchange of computational tools.

Keywords Nitinol · Shape memory alloys · Modeling · Data-driven

Introduction

Data Science—a collection of scientific methods to gather and analyze data—has emerged as a versatile tool to advance the descriptive and predictive capabilities in various areas of science and engineering. As an example of the descriptive capabilities, data science enables determining parameters that most influence the response of a system. It also enables quantification of uncertainty in the measurements of the behavior of a system. In terms of the predictive capabilities, data science enables the development of predictive models for the response based on the observed data and not on analytically derived functions based on a postulated form or ansatz. The tools in this data-driven paradigm such as high-speed and multi-modal data acquisition, efficient data storage and retrieval, and statistics-related learning techniques such as machine learning (ML) have already found several applications specific to the mechanics of structural materials [1, 2]. Cardiovascular implant design that relies on the mechanics of the underlying materials can greatly benefit by adapting this data-driven paradigm.

Nitinol or Nickel–Titanium shape memory alloy, which is one of the most common materials used in these devices, exhibits monotonic and cyclic mechanical behavior that is complex enough to warrant over four decades of continuous scientific attention. The monotonic mechanical response of Nitinol is challenging to model and predict due to the presence of superelasticity, anisotropy, and strong dependence on the processing parameters [3]. The cyclic response of Nitinol is even more challenging to predict

This invited article is part of a special topical focus in the journal Shape Memory and Superelasticity on Cardiovascular Nitinol Medical Devices. The issue was organized by Dr. Srinidhi Nagaraja, G.RAU, Inc. and Dr. Harshad M. Paranjape, Confluent Medical Technologies, Inc.

✉ Harshad M. Paranjape
harshad.paranjape@confluentmedical.com

[1] Confluent Medical Technologies, Inc., 47533 Westinghouse Drive, Fremont, CA 94539, USA

[2] Department of Materials Science and Engineering, The Ohio State University, 140 W. 19th Avenue, Columbus, OH 43210, USA

Published online: 08 February 2023

because of the structural and functional fatigue caused by the interaction between the deformation mechanisms of phase transformation and plasticity, the complex role played by non-metallic inclusions in initiating a fatigue fracture, and the large-cycle regime (10^7–10^9 cycles) relevant to the cardiovascular implants [4, 5].

The phenomena mentioned above are made possible by a diffusionless phase transformation that can reversibly occur in Nitinol in response to an applied stress or a temperature change [3, 6]. The transformation occurs between two phases referred to as austenite and martensite. The austenite phase in Nitinol has a cubic crystal structure. The martensite phase mainly has a monoclinic crystal structure, but it can have a trigonal crystal structure as well. In the latter case, the martensite is referred to as the R-phase. The stress-induced, reversible martensitic transformation to the monoclinic martensite phase results in a large recoverable strain and furnishes the phenomenon of superelasticity that is widely used in implants. This property allows a Nitinol stent to be crimped to a compact size and deployed to a larger size without any permanent deformation. The microstructure of the phase-transformed Nitinol is however, quite complex. It consists of domains of austenite, domains of monoclinic martensite with its own nanostructure called twinned martensite, in some cases domains of R-phase, and regions of plastically deformed austenite and martensite. This is in addition to the microstructural features that are present in all alloys such as grains and subgrains formed during the solidification and heat treatment of alloys, non-metallic inclusions due to oxygen, carbon and other impurities inherited from air or vessels used in the solidification of the alloys, and dislocation networks due to the cold work imparted during the processing of alloys. During the use of an implant made using Nitinol, the microstructure evolves due to nuanced processes such as the transformation-induced slip around growing and shrinking martensite domains [7] and large stress concentrations around non-metallic inclusions [8] that can cause local plastic deformation.

Because of these complexities, the simulation of Nitinol-based implant deformation typically relies on phenomenological models that rarely account for the specificities of pre-processing and microstructure that affect the mechanical behavior of Nitinol materials used in individual products. The macro-scale, phenomenological models that are typically used in the device-scale simulations rely on small-strain formulation, assume isotropy, and opt to mostly not capture the coupling between phase transformation and plasticity. One of the more common Nitinol simulation implementation used in the industry is based on the work of Auricchio and others [9]. This model is generally able to capture the superelastic response and the asymmetry in tension and compression response. Additions to the implementation in commercial finite element analysis tools such as Abaqus and Ansys now allow simulation of plastic deformation, hardening of the transformation response, and evolution in the transformation behavior with cyclic loading. More recent advancements to the simulation of Nitinol behavior are reviewed by Cisse et al. [10]. The macro-scale models, relevant to the device-scale simulations, are either based on phenomenology or energetics of inelastic deformation. In either case, the form of a model depends on the aspects of the deformation considered by the developer. For example, the developer may choose to model the multi-step austenite $\leftrightarrow$ R-phase $\leftrightarrow$ martensite transformation but choose not to model the reorientation of martensite variants. Zaki, Moumni, and colleagues have developed a series pf phenomenological models that capture aspects of Nitinol deformation such as reorientation of self-accommodated martensite, superelastic deformation, plastic deformation, coupling between phase transformation and plasticity, tension-compression asymmetry, and evolution of response on cycling [11–13]. These models do not capture multi-step phase transformation and functional fatigue. The model of Peultier et al. is based on the energetics of a representative volume element consisting of an aggregate of austenite grains and martensite domains [14]. They are able to capture both superelastic and shape memory response. This model does not capture most other aspects of deformation exhibited by Nitinol. Thus, even though a variety of models exist for simulating the deformation of Nitinol, not a single model can be guaranteed to capture the response of an implant manufactured using a particular form of Nitinol and having undergone a specific sequence of processing steps. Additionally, it is common for complex models to require significant effort to determine the input material properties and they tend to be slower. The versatile data-driven paradigm holds the potential to advance the simulation of Nitinol deformation by providing two key benefits: a constitutive law that is directly built on the observed response and not based on a limited set of selected deformation behaviors and a potential speedup due to the data-based learning methods used in the implementation of these simulation methods. Considering how commonplace computer simulation is in the design of Nitinol-based cardiovascular implants, this advance can positively impact the industry practice.

Since Nitinol deformation simulation is expected to be predictive, it is essential to develop reliable methods to acquire data that will act as inputs to the simulation. Experimental data are used in simulations of Nitinol deformation for various purposes. Most fundamentally, the experimentally measured constitutive response data or the stress–strain curves inform the deformation modes that need to be simulated. Tensile testing data on medical-grade

Nitinol inform us that elasticity, superelasticity, and plasticity are three deformation modes through which Nitinol deforms. Experimentally measured constitutive response data are also used to determine simulation model parameters. This process is typically known as model calibration or material property determination. Finally, experimental data are used to validate the simulation results. For example, the superelastic material properties in many medical device deformation simulations are determined using tensile test data. Then, the simulation results are validated by performing a separate simulation of a different deformation mode such as bending or radial loading and comparing the results with corresponding experimental data.

The experimental data acquired for these purposes could be global, full-field, or mixed-mode. Global data refers to measurements such as the load measured on a specimen that is tested in tension or the tensile strain averaged over the gage of the specimen. Full-field data on the other hand provides a spatially resolved measure of deformation. Full-field measurements provide a larger quantity of data in a single measurement compared to global measurements. Thus, full-field deformation measurements have become commonplace for informing simulations or for validation of the models. However, in many instances, the availability of input data are limited or it is time consuming to acquire experimental data from a particular source. Thus, it is advantageous to merge data from multiple available sources. Above all, simulation is one of the available tools when designing Nitinol-based medical devices or when demonstrating durability of such devices under a particular use case. Because of this, it is essential to quantify and communicate the credibility of data used to build the simulations and also quantify the credibility of simulation results [15, 16]. Communication of simulation credibility helps the stakeholders in determining the extent to which they can rely on the simulation results to make decisions regarding device safety and efficacy. Quantification of credibility using methods generally accepted by the community helps communicate the simulation credibility objectively. In other words, with the knowledge of the simulation credibility, the role played by simulations in a risk-informed decision making scheme can be appropriately determined. For example, simulations used for durability assessment can be of greater utility if a simulated fatigue safety factor and the uncertainty associated with it are both known.

With this motivation, we review the recent advances in the field and provide a perspective on the adaption of the data-driven experimental and modeling methodologies to Nitinol mechanics. Considering that the field is nascent, we propose a three-tiered approach suitable for a gradual adaption of these methodologies. The aim behind proposing this approach is to provide an organization for the body of relatively new methods reviewed in this work. Hence, the tiers in this approach are an example and practitioners can adapt them to suit their resource availability and target timeframe. The tiers are illustrated in Fig. 1. The first tier relates to the acquisition of Nitinol material property data and quantification of the uncertainty in the data. This is referred to as the *information* tier. The second-tier relates to fusing multi-modal data to reveal mechanisms that fundamentally determine the Nitinol mechanical behavior. This is referred to as the *knowledge* tier. The third tier relates to making performance or property predictions-based directly on the available data rather than relying on subjectively derived analytical expressions or phenomenological models based on a postulated form or ansatz. This is referred to as the *prediction* tier.

Recent advances in the data-based methods organized in these three tiers are summarized in the next three sections. We first describe methods in information tier that enable acquisition of Nitinol material deformation data and also enable quantification of the credibility of that data. Then we describe methods in the knowledge tier that facilitate assimilation of data from multiple sources to uncover knowledge about deformation mechanisms. Finally, we review methods in the prediction tier to simulate the deformation of elasto-plastic materials. We provide a perspective to extend these methods to simulate the deformation response of superelastic Nitinol implant components. These methods fundamentally differ from the existing simulation approaches in that they learn the stress–strain relationship or the constitutive law based on any available deformation data and depend less on developer-chosen postulates regarding the form of the constitutive law. We close by summarizing the review, providing a specific idea of Nitinol simulation where some of these methods can be applied, summarizing the limitations of this approach, and providing a short discussion on promoting a broader adaption of these data-driven methodologies by the Nitinol community.

The adaption of methods described here demands a substantial investment in terms of resources required for their research and implementation. The gains from adaption of these methods will be incremental. But these methods overcome the following key challenges related to the current methods for characterization and simulation of Nitinol implants:

1. Uncertainty quantification is an integral part of these data-driven methods. Presentation of simulation results and uncertainty together makes these methods more suitable for serving as in-silico evidence for demonstrating durability of Nitinol implants. Verification,

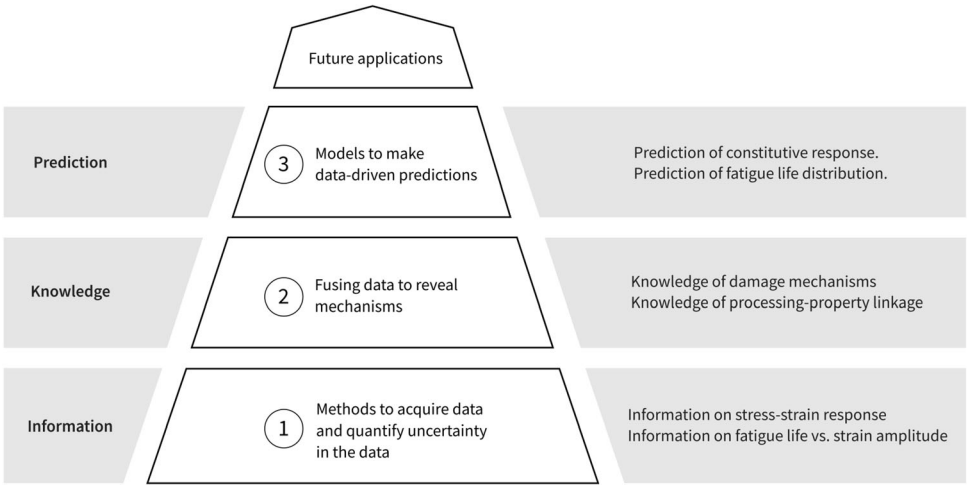

Fig. 1 An example three-tiered approach to incorporating data-driven methods in the modeling and simulation of Nitinol mechanics

validation, and uncertainty quantification with existing simulation methods is generally an afterthought and it often requires comparable or larger effort.

2. The data-driven simulation approach requires a relatively large upfront development effort but reduces the need for ad-hoc changes to the simulation framework in order to simulate phenomena specific to a Nitinol raw material or a device. For example, established simulation models for superelasticity can be augmented with some effort to enable simulation of R-phase formation. But a model-free, data-driven simulation approach can intrinsically enable the incorporation of R-phase transformation in the simulated response if the model is trained on data that exhibits the R-phase transformation.
3. The data-driven constitutive modeling methods promoted in this work use computational learning methods that are intrinsically designed to use modern computational hardware such as graphics processing unit (GPU) to accelerate computation. Existing computational methods have been relatively slow to enable the use of these modern computational hardware. Thus, the new approach is more suitable to benefit from the rapid advancement in computer hardware including the use of hardware specifically designed to run the implementations of these learning techniques.

With these key benefits of the new data-based simulation paradigm for Nitinol implants, we are optimistic about the adaption of these methods through example tiers described below.

Methods to Acquire Data and Quantify Uncertainty in the Data

Full-Field Data Acquisition Methods

Several methods have been demonstrated for measuring the full-field deformation of metals and specifically, the deformation of shape memory alloys [17]. These methods include surface measurement techniques such as digital image correlation (DIC) and Moiré Interferometry and volume measurement techniques such as 3D high-energy X-ray diffraction [18, 19].

The DIC technique for measuring full-field surface strain field has been widely adapted in academic settings [20]. DIC relies on imaging a quasi-random pattern on the surface of the test specimen at certain intervals and analyzing the change in the pattern to calculate surface displacement and strain fields [21]. In relation to mechanical behavior simulation, an important use of DIC is to identify the simulation inputs. Most common simulation inputs determined using DIC are the material properties. This process is often referred to as model calibration. Pierron and co-authors provide a detailed overview of various methods for identifying the material properties using DIC measurements [22, 23]. The large amount of strain data acquired using DIC can help obtain more accurate material properties for complex constitutive models such as those for Nitinol superelasticity. Another application of DIC is in the validation of finite element analysis (FEA) models of deformation [24]. Comparing local deformation results between simulation and DIC measurement can provide a more appropriate validation of the model versus comparing just the global results such as load. While the DIC technique has received wide adaption in the automotive, aerospace, and other industries, it has received relatively small adaption in the medical device design industry.

Aycock and co-authors recently described a detailed framework for measuring full-field deformation on the surface of Nitinol medical devices [25]. They describe the complete pipeline for acquisition of full-field strain data including the details of a miniature load frame appropriate for loading small-scale Nitinol implants, the method for acquiring DIC data, and a snapshot of representative results for the strain data. Quantification of error in the measured strain data are useful in quantifying the credibility of the DIC data. In that regard, they also provided detailed methodology and analysis for quantifying the error and the noise in the acquired data. Acquisition of full-field deformation data from actual medical devices can provide a more effective means of validating computer simulation models used in the design or durability assessment of that particular device. Senol and co-authors have described the use of DIC to validate fatigue strain simulations of a Nitinol test specimen [26]. They demonstrate the use of 2D DIC for measuring mean and alternating strains during cyclic loading of a Nitinol surrogate specimen. Their results demonstrate that the strain amplitudes compare well between DIC measurements and simulations, but the mean strains differ. The difference is particularly notable in the medium mean strain values where phase transformation occurs through a part of the gage of the specimen. These approaches can be combined to develop an end-to-end workflow to inform and validate the simulations used in the durability assessment of a Nitinol device. Such a workflow can even provide an experimental substitute to the calculation of fatigue safety factor of a Nitinol implant subjected to particular boundary conditions. An example workflow for the direct evaluation of fatigue indicator parameters and potentially for the evaluation of the fatigue safety factor is shown in Fig. 2 and further described in [27]. The workflow first establishes a standard test specimen representative of the geometry of that particular medical device (Fig. 2a). Here we show a diamond specimen that has historically been used to represent the unit cell of a typical stent. However, the workflow is generic and can be applied to other specimen geometries as well [28, 29] and can be applied to test coupons cut from an actual device. Then, a DIC test setup to measure the surface strains in the specimen is constructed (Fig. 2b). A variety of commercial DIC setups are available or they can be built using an appropriate selection of cameras, lighting systems, environmental control solutions, and DIC post-processing software. A cyclic loading protocol (Fig. 2c) to impose a cyclic deformation on the specimen is needed. While the

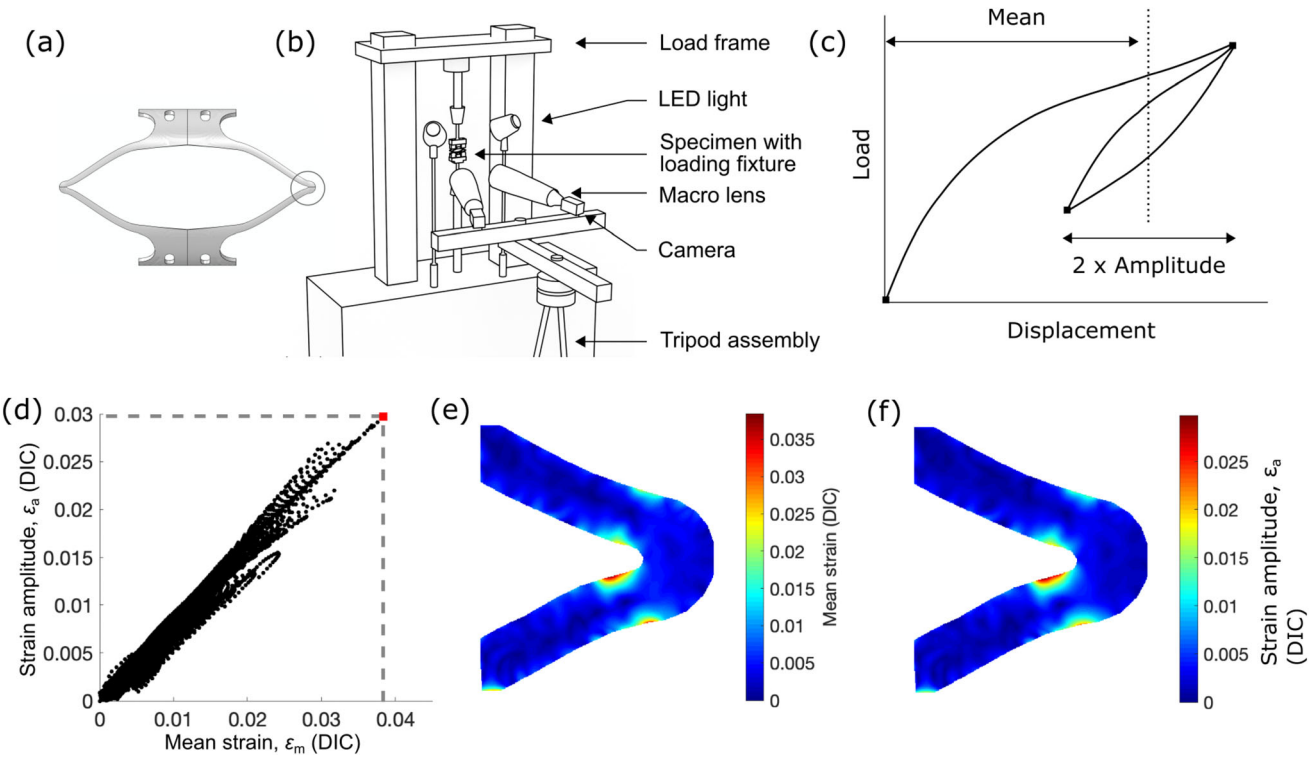

Fig. 2 Full-field strain measurement on a diamond Nitinol test specimen during cyclic loading. **a** Schematic of the diamond specimen. **b** Schematic of the digital image correlation (DIC) setup. **c** Schematic of the global load–displacement response showing initial monotonic loading and then cyclic loading in a subcycle. **d** Strain point cloud in the subcycle obtained from full-field surface strains on the diamond measured using digital image correlation. **e** Experimentally measured mean strain distribution in the diamond apex region. **f** Corresponding strain amplitude distribution

protocol can be as simple as a pre-deformation followed by cyclic application of a displacement range, a protocol that is more representative of the anatomical boundary conditions experienced by the implant can be developed. As the deformation of Nitinol is dependent on strain rate and temperature, it is necessary that the test protocol is representative of these deformation conditions that an implant will be subjected to. This may require testing of the material in an aqueous medium and at a quasi-static strain rate. Commercial DIC implementations typically enable this, though the experimental setup could become cumbersome due to these requirements. Using these three pieces of infrastructure, the surface strain fields on the implant can be measured. The surface strain fields can be used to calculate the fatigue strain map or the point cloud of strain amplitude vs. mean strain as shown in Fig. 2d using an appropriate scalar or tensor method [30]. The data in turn can be used to plot distributions of the mean strain (Fig. 2e) or the strain amplitude (Fig. 2f) on the device profile. These data can be subsequently used to determine the fatigue safety factor. While this approach is limited to the fatigue strain assessment on the specimen surface and perhaps one surface of the specimen geometry facing the DIC camera system, it may be adequate since critical fatigue strains often occur on the specimen surface rather than in the interior volume for many common device geometries [27].

As described in the example above, the DIC technique has tremendous potential utility in the quantification of medical device deformation and in complementing simulation. Continuing advances in the test hardware will make this technique more accessible to the medical device industry. Advances in data science are also leading to advances in the full-field deformation characterization methods. Recently several authors have implemented algorithms for full-field deformation quantification using deep learning, a statistical data analysis method in the family of ML techniques. For example, Boukhtache et al. developed a convolutional neural network framework that can be trained on a specific speckle pattern and strain dataset and used to obtain displacement and strain fields from other speckled specimen images [31]. This approach can provide subpixel resolution at a speed comparable to DIC when displacements per step are less than one pixel. Similarly Yang et al. have developed a framework of two convolutional neural networks—DisplacementNet and StrainNet—for DIC-like full-field deformation quantification [32]. A key contribution of their approach is that the StrainNet neural networks can directly predict strain fields without having to determine displacement fields as an intermediate and then calculating numerical derivatives to obtain strains. They demonstrate that this approach is able to furnish superior results to some of the commercial DIC offerings at large strains and in large localized strain conditions. We note that these techniques provide an alternative to DIC for obtaining surface strain fields and currently they are limited to 2D strain fields and do not include depth sensing. Zhu and co-authors have applied computer vision techniques to obtain surface strains [33]. Their approach differs from the two deep learning methods described above and also fundamentally differs from DIC in that, Zhu et al. implement a multi-scale image feature registration algorithm that progressively determines the deformation between a reference image and a result image by sampling those images at various resolutions. Both DIC and deep learning methods on the other hand analyze images at a certain subset level and furnish strains at that same length scale. These advancements mean that full-field data acquisition can be a useful new tool for informing and validating medical device simulations.

Methods for Uncertainty Quantification

While inputs, methods, and outputs of medical device simulations often receive extensive scrutiny, one aspect that generally receives a lower attention is the uncertainty of the inputs and its impact on the credibility of simulation outputs. For example, in simulation studies for assessing fatigue safety of medical devices, it is typically seen as adequate to report a fatigue safety factor greater than one. In reality, a fatigue safety factor greater than one means little if the simulation is based on inputs that have a large uncertainty associated with them. Thus, reporting of uncertainty in experimental or simulation data are just as important as acquiring that data.

Recently, various efforts have developed consensus methods of evaluating and communicating credibility of computational models [15, 34, 35]. Overall, they recommend that the credibility of a computational model should be commensurate with the level to which its results are used in making final decisions related to device efficacy or safety. Once an expectation on the credibility of computational model is set, some of the methods reviewed below can be used to assess the uncertainty in the inputs to the computational models and then propagate those input uncertainties to the simulation results.

Ricciardi and co-authors propose a framework for quantifying uncertainties in simulation model parameters such as the hardening parameters for a plastic deformation model when the parameters are determined using experimental data [36]. Their approach, based on Bayesian inferential framework, can be used to predict the uncertainty in the mechanical behavior of samples that are not yet tested. They provide a detailed discussion on the formulation of a Bayesian model for uncertainty quantification and propagation of uncertainty to subsequent

simulations of deformation. The Bayesian approach is a statistical method that allows estimation of the probability of a hypothesis being correct based on the available information. This is an appropriate tool for structural mechanics when the intent is to determine the probability distribution of certain mechanical parameters based on available experimental data. The use of Bayesian inference applied to model parameter determination is a well-established practice in structural mechanics [37, 38] and other engineering disciplines [39]. However, until recently, this method has not been applied to the simulation of medical device deformation. Paranjape and co-authors recently applied the technique to determine the probability distribution of superelastic FEA material parameters determined from experimental full-field strain and global load data [40]. They also demonstrated that the uncertainties in the superelastic material parameters can be propagated to the simulations of fatigue safety factor determination. This furnishes a fatigue safety factor value and its credibility intervals. A narrow credibility interval corresponds to a lower uncertainty in the results. Their approach consists of developing a ML surrogate model for the superelastic FEA model so that the sampling of the material property distribution can be performed more efficiently.

The topic of using probabilistic methods such as the Bayesian method for calibrating the inputs of a constitutive model receives some resistance because of the perception that they incur significantly higher upfront effort compared to the more typical ad-hoc methods. There continues to be development of more efficient methods such as efficient sampling strategies for Bayesian calibration that will continue to reduce this burden. Here sampling refers to the numerical methods that are used to obtain the full probability distribution of a parameter under the framework of Bayesian inference. Sampling can be a computationally burdensome effort. lye and co-authors provide a tutorial on various sampling methods in the family of Monte Carlo methods to illustrate the rate of convergence, numerical challenges, and time required to obtain a probability distribution of a parameter [41]. More generally, Cranmer and co-authors list opportunities for advancement for these probabilistic methods in three areas - better determination of uncertainty distribution using smaller input datasets, improving how accurately the uncertainty distribution is determined based on the input model parameters, and making uncertainty quantification more modular such that new data can be sequentially plugged in to update the probability distribution [42].

The statistical methods to quantify the uncertainty in simulation parameters and results are well-established and implemented in a variety of common computational tools such as Python and Matlab. As reviewed above, full-field data acquisition methods that can be used to inform such Bayesian inference methods of input parameter determination and uncertainty quantification are also well-established. Thus, we are in a position to broadly adapt these methods and make input and output uncertainty quantification a standard practice in the simulation of Nitinol medical device deformation. The methods described here can also be applied to other simulation inputs such as anatomical boundary conditions that are determined from computed tomography, radiography, or other experimental means.

The methods described in this section were in the information tier of the three-tiered approach for adapting data-driven methods to Nitinol implant design and simulation. The knowledge tier methods described below to uncover patterns and mechanisms for deformation phenomena such as fatigue benefit from the methods described in this section.

Methods to Fuse Data and Uncover Mechanisms

Many mechanistic problems encountered in the design of cardiovascular implants cannot be directly solved by acquiring experimental data using the sophisticated methods described above. Nor they can be solved by sophisticated simulations alone. Modeling of fatigue is one such problem. Fatigue is a loss in functionality, loss in strength, or catastrophic fracture in components due to cyclic loading. An implantable medical device is subjected to cyclic loading due to the cardiac rhythm. If such a device is improperly designed, it can lead to a fatigue failure. Thus, demonstration of durability is an important step in the design of implants. A key input in determining the durability of an implant is the fatigue resistance of the underlying Nitinol material itself. It is generally accepted that the fatigue resistance characterization of the base material should be performed to an equivalent number of cycles as the actual device is expected to be exposed to [43]. If the implant under consideration is in the structural heart space, this means that the base material might need to be tested to 600 million cycles. Such testing is expensive and time consuming. Instead, if a model is available that can predict the fatigue resistance distribution at a given fatigue life or the fatigue life distribution at a given resistance as a function of microstructural and macro inputs, it can significantly reduce the amount of testing required. Further, such model can be used to estimate fatigue resistance even if changes to material purity and pre-processing factors such as pre-straining or cold work are made at an intermediate phase during the implant design process. Thus, we first review data fusion approaches in the literature related to fatigue life prediction of elasto-plastic and superelastic

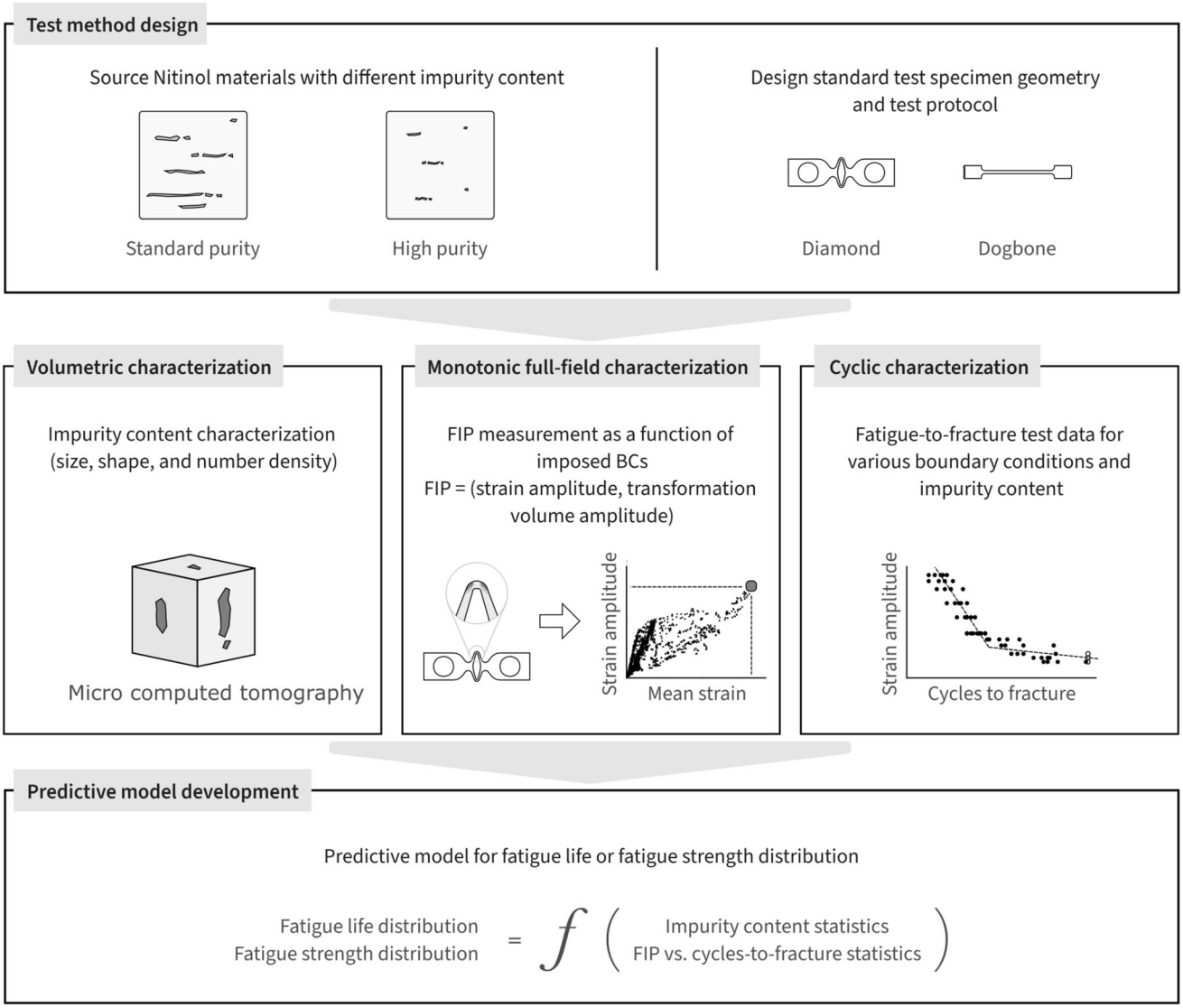

Fig. 3 An example of a data fusion strategy for developing a predictive model for fatigue life or fatigue resistance of a Nitinol base material. FIP stands for fatigue indicator parameter

materials and particularly focus on approaches that have used data science tools such ML.

Sangid reviewed opportunities for combining data from multiple sources to more efficiently predict deformation and failure modes of structural materials [44]. A particular opportunity they mention is using microstructural simulations to augment the fatigue data so that fatigue strain-life curves can essentially be extrapolated to lower fracture probabilities. Chen and Liu provide a comprehensive review of ML approaches used in the modeling of various fatigue-related phenomena [45]. In one particular example, Durmaz et al. describe a modular and automated workflow to test fatigue specimens, measure specimen deformation, automatically identify damage initiation regions, and develop a predictive model for fatigue initiation sites in metallic materials [46]. While their work centered more on microstructural features such as grain boundaries, materials such as stainless steel, and fatigue crack growth monitoring rather than on fatigue life itself; the multimodal fatigue and microstructure data acquisition workflow described by them is worth adapting for Nitinol fatigue life prediction. Gebhardt et al. combined impurity shape information from micrographs and fatigue strength information from microstructural simulations to develop a model for fatigue life of nodular cast Iron [47]. The impurity size and shape distribution were the inputs to the model. They used a ML tool called simplified ResNet. These impurity or inclusion size distributions have received wide attention in many alloy industries, because inclusions are perhaps the most important microstructural entities relevant to fatigue

[48, 49]. For example, Murakami and coworkers have developed models for fatigue strength as a power-law function of the inclusion area fraction [50, 51]. Some work has been done on describing the size distribution of inclusions using methods such as extreme value statistics [52] and on quantifying their effect on the fatigue response of Nitinol [53]. Despite this rich literature on the connection between metal fatigue and inclusion size statistics, limited work has been done using learning methods such as ML to predict fatigue life as a function of the inclusion statistics. Specific to Nitinol, Kafka and co-authors, building on the work of Moore et al., developed an approach to predict the fatigue life of Nitinol at a fixed strain amplitude as a function of particle-void-assembly size and location [54, 55]. They combine highly simplified particle or inclusion size data from 3D characterization with a microstructural model for plastic deformation. Their results capture the generally understood trend that the fatigue life decreases with increasing impurity particle size. They also predict that the fatigue life increases with increasing void size for a void that is present between adjacent inclusion particles. Inspired by these examples and particularly by the work of Durmaz et al. on prediction of fatigue life based on microstructural data from multiple sources, we now provide a perspective on an approach to predict the fatigue life of Nitinol based on microstructural information.

An example approach based on data fusion that can be used to predict fatigue life or fatigue resistance of Nitinol base materials is illustrated in Fig. 3. The goal of this approach is to predict the fatigue life distribution or fatigue strength distribution of a Nitinol base material when information on the impurity or non-metallic inclusion shape and size distribution is known. By *Nitinol base material* we mean sheet, strip, tubing or similar Nitinol material forms that have undergone a particular processing sequence. By *fatigue life* we mean the number of cycles required to cause fracture at a given strain amplitude. By *fatigue resistance* we mean the strain amplitude requited to cause fracture at a given number of cycles. By *distribution* we mean a median value and associated uncertainty bounds or an equivalent uncertainty quantification. We propose that the proposed approach is developed in three phases. The three phases, schematically shown in Fig. 3, are described below.

The first phase, named the *test method design* phase, involves selection of Nitinol materials of various microstructural attributes and the design of a test specimen that can be used to make a fatigue test coupon. This may include materials with differing Oxygen and Carbon content, particle-void-assembly area fraction, inclusion morphology distribution, and cold work. Diamond, dogbone, and C-shaped dogbone [28] are some of the specimen geometries that may be considered for making the test specimens from these materials.

The second phase—the *characterization* phase—involves a multi-modal characterization of the microstructure and the fatigue performance of these materials after they have undergone appropriate processing such as pre-deformation. The microstructural attributes that most significantly influence the fatigue performance will be selected. This selection can be performed using statistical methods such as principal component analysis or it can be informed by data in the literature regarding property-fatigue performance correlation. Significant work on identifying key microstructural parameters influencing fatigue has been performed on Ti-based alloy systems [56, 57] and it can be extended to Nitinol. These microstructural attributes will be characterized in the Nitinol materials using 2D and 3D methods. 2D characterization methods include metallography and scanning electron microscopy to identify non-metallic inclusion fraction. 3D methods include micro-computed tomography that provide volumetric information on the microstructural elements. A full 3D characterization of the microstructural elements may be more desirable because of the anisotropy inherent to most drawn Nitinol products. A monotonic and cyclic performance characterization will be performed on the test specimens manufactured from these materials using displacement boundary conditions that result in a range of mean strains and strain amplitudes that typically occur in cardiovascular implants. The monotonic characterization will be performed to determine the relation between displacement boundary conditions and strains. While this function is typically performed using simulation methods such as FEA, a fully experimental approach using methods such as DIC, as described in Fig. 2 may be advantageous. Cyclic characterization will be performed under these displacement boundary conditions to obtain fatigue-to-fracture data. These data can be later used to determine fatigue resistance for a specific fatigue indicator parameter such as the strain amplitude, transformation volume amplitude, or a stress- or energy-based fatigue indicator. Fatigue-to-fracture data allows calculation of the fatigue strength distribution or fatigue life distribution. The testing will be performed to a reasonable number of cycles representative of the clinical area where these models will be eventually used. For example, if the fatigue data will be used for Nitinol heart valve design, then tests to 600 million cycles should be performed. If the data will be used for the design of peripheral vascular prostheses, then tests to a lower cycle count such as 10 million cycles could be performed. The monotonic and cyclic characterization together provides a map of the fatigue resistance of the material as a function of mean strain and strain amplitude or other similar fatigue indicator parameters. While a damage tolerant approach is

less commonly used in the design of durable Nitinol implants, fracture mechanics and fractography may provide additional insights on the fatigue failure mechanisms.

In the third phase, referred to as the *modeling* phase, a model for fatigue resistance and fatigue life will be constructed. The fatigue resistance and fatigue life will be modeled to be a function of the microstructural parameters such as impurity size and a function of the fatigue indicator parameters such as mean strain and strain amplitude. Relatively simple regression methods can be used to construct the model. However, recently, ML and other statistical methods have made several other tools available for such model construction [45]. Many prior works in the literature related to fatigue life model generation have focussed on the stress-life models. However, the methodologies documented in those works can be extended to strain-life approach relevant to Nitinol. Barbosa and co-authors described a ML method to fit stress-life models to mean stress, stress amplitude, geometric parameters, and fatigue life data [58]. The inputs to their ML model are the mean stress and a target fatigue life. The ML model, based on a shallow neural network, furnishes the stress amplitude corresponding to the inputs. Chen and Liu developed an ML model for determining fatigue life and the uncertainty in the fatigue life as a function of multiple inputs [59]. Their shallow neural network-based model is able to incorporate mechanistic constraints based on common observations related to the fatigue life vs. mean stress relationship. These constraints could include the presence of a *knee* seen in the stress/strain-life curves of some materials. Their model is also able to incorporate both fracture and runout data. Statistical methods other than ML can also be used to develop models of fatigue life. Dourado et al. developed a Bayesian fitting method for strain-life data [60]. Their method enables the quantification of uncertainty in the fatigue life predictions performed using the model. The fatigue life estimation itself is assisted by a Coffin-Manson-type law and they assume a log-normal distribution for the fatigue life. Whichever is the method used to construct the model, the utility of this data-based approach is that the fatigue life or fatigue resistance of materials that have not been tested before can be predicted using this data-driven framework. The fatigue model development described here may also be able to uncover new dependencies between fatigue life and various processing parameters or microstructural parameters.

The data fusion approach described here for predicting fatigue resistance aids in uncovering fatigue mechanisms by providing a holistic view of the cyclic deformation response of Nitinol. But, a holistic view by itself does not directly reveal fatigue mechanisms. It can reveal material parameters that primarily influence the fatigue behavior or reveal non-intuitive dependencies. However, we may still have to make interventions when uncovering mechanisms using the data and resources that are available to us. We will also have to be creative in designing experiments that capture nuanced effects of microstructure on fatigue such as the effect of stress field around inclusions on fatigue life. As the data acquisition methods and fatigue model development methods become more efficient and robust, a progressively reduced oversight may be possible in the development of predictive models for fatigue and we will be able to capture the effect of more complex aspects of microstructure on fatigue.

The methods in the knowledge tier, described in this section, are primarily descriptive and in part predictive. We now provide a review and perspective on predictive data science methods for Nitinol design and simulation.

Methods to Make Data-Driven Predictions of Nitinol Deformation Response

The most impactful utility of data-driven methods is in terms of their predictive capabilities. The ability to predict the deformation response of a Nitinol component under various boundary conditions on the basis of a limited information on the constitutive response of the base material is essential in the design of Nitinol implants. Various constitutive modeling and simulation tools such as FEA typically provide this capability. Constitutive response of a material is the relation between stress and strain. Knowledge of accurate constitutive response is one of the most important prerequisites for reliable simulation of Nitinol implant deformation. While the implementation of the constitutive response of Nitinol in simulation tools has incrementally advanced over the years, a leap in realistic constitutive modeling can significantly enhance the accuracy of simulation of implants. It will also contribute to making in-silico evidence of device durability a widely accepted part of the device regulatory approval process. A field where data science methods are used to develop simulation methods for the constitutive response has emerged. These methods are often referred to as data-driven constitutive modeling. Several methods have been proposed for the data-driven simulation of elasto-plastic deformation. If applied to the superelastic deformation simulation, these methods can significantly advance the simulation of Nitinol deformation.

Data-driven methods of constitutive response simulation can be broadly grouped in two closely related categories: Surrogate models and model-free data-driven methods [61]. Surrogate models use stress–strain data from existing state-of-the-art simulation methods and create a computationally efficient surrogate regression model using methods such as deep neural networks. Regression-methods-based

surrogate models can be efficient because they do not need to iteratively solve the field evolution equations in order to determine the local load or stress response. Model-free data-driven methods on the other hand train the constitutive response based on indirect data. For example, they derive the stress–strain response at a material point based on an experimental dataset consisting of global load and local surface strain histories. This derivation of the constitutive response is generally performed under the constraint of applicable physical laws such as equilibrium and conservation of energy. Model-free methods are notable such that they directly derive the deformation mechanics from the experimental data and do not rely on expert judgement or local stress–strain history from a pre-built higher-fidelity simulation data library to define a particular model form.

Development of surrogate models for history-dependent constitutive response such as elasto-plastic response has been explored for several decades [62]. However, it has become feasible in the last few years due to the increased availability of computational resources and the availability of easy-to-use implementations of sequence-learning methods such as long short term memory (LSTM) and gated recurrent units (GRU). Work of Mozaffer et al. was one of the first studies to develop a surrogate model for path-dependent plastic deformation response in 2D [63]. They used a simulation library as an input and used the GRU sequence-learning ML method. Their method is able to capture material hardening and stress concentration due to local inhomogeneities. Other similar approaches have demonstrated enhanced capabilities such as modeling of constitutive response with anisotropic yield behavior [64], using temporal convolutional network (TCN)—another ML sequence-learning technique—to provide a surrogate for visoplastic and temperature-dependent response [65], using Linearized Minimal State Cell—a form of recurrent neural network—to obtain a data-driven plasticity surrogate model from long sequences of stress–strain history data [66], and the use of internal state variables [67]. Huang and co-authors propose a method for building a data-driven history-dependent inelastic constitutive law where they suggest obtaining the stress–strain sequence data for training from experiments on specialized specimens such as biaxial cruciform geometry [68]. Designing specialized specimen geometries is essential for obtaining strain paths that cover a large portion of the strain space such as tension, compression, non-proportional loading, etc. Liu et al. present another innovation in developing a surrogate model for plasticity based on the simulation data obtained at a finer length scale [69]. That is, they train a data-driven surrogate model on the stress–strain data obtained using finer-length-scale simulations that implement constitutive modeling techniques such as crystal plasticity and then use the surrogate model in the larger-length-scale simulations that yield macro-scale stress as a function of a macro-scale strain increment. The work of Karapiperis et al. is in the same domain as it seeks to build a data-driven constitutive law based on lower length-scale simulation results [70].

Model-free data-driven methods of simulation for inelastic materials have been developed in the last few years. Eggersmann and co-authors developed a data-based approach to build a constitutive law [71]. Their approach is able to incorporate history-dependence by directly training the data-driven constitutive law on the strain history or through incorporation of internal variables. While their approach fundamentally depends on the availability of local stress–strain data pairs to train the constitutive law, their collaborators have demonstrated methods that can extract the requisite local stress–strain data from the macro-scale boundary conditions and local strain fields [72, 73]. Since these approaches provide options to incorporate internal variables or history variables in the constitutive law determination, they could be appealing for extension to superelasticity simulation where martensite phase fraction can act as an internal variable to capture the history dependence of deformation. The recent work of Langlois et al. provides an extension of this approach where the constitutive law is obtained from an initial approximation of the local stress state [74]. The approximate local stress state is obtained from the local strain field using the finite element model updating (FEMU) method. The local strain field can be measured using methods such has DIC. They demonstrated the method on an elasto-plastic material showing strain localization and thus, this method could lend well to the simulation of superelastic response that occasionally exhibits strain localization features such as Lüder's bands. Cameron and Tasan [75] propose an approach that can extract full-field stresses from full-field strain and traction boundary conditions. Their approach requires the knowledge of elastic strain component at each point, which could be challenging. However, this approach is still feasible if the material behavior is assumed to be elastic and the elastic modulus is known from an initial analysis of the experimental stress–strain response. Once the full-field stresses and strains are known, a constitutive law can be constructed using one of the sequence-learning methods cited in the paragraph above. Flaschel et al. propose an approach that obtains a constitutive law with an ansatz selected from a catalogue of predefined functions [76, 77]. With a reasonably large prebuilt catalogue, this approach can be suitable for modeling any complex stress–strain response.

Each method in this broad collection of data-driven methods has sought to address specific nuances of constitutive law development. Yet all of these methods rely on certain non-trivial statistical techniques such as

minimization or ML-based regression to obtain the constitutive law. Thus, these methods tend to be critiqued on two common aspects. First, there are concerns that the *black-box* statistical formulations used in many of these techniques make it challenging to check for any violations of the fundamental thermodynamics or statics principles. This is particularly critical when the data used for training the constitutive law is noisy and contains outliers, which if taken at their face value can lead to non-physical deformation modes. Some efforts such as the work of Masi et al. have attempted to address this concern by encoding the fundamental thermodynamic conservation principles in the structure of the constitutive modeling ML framework itself [78]. This approach is part of a broader effort to develop physics-informed ML methods for various physics and mechanics problems [79]. Second, there are concerns that the large amount of data used to train the constitutive laws using these methods makes it challenging to quantify the uncertainty in the predictions from these models. Thus, there have been attempts to quantify the uncertainty in the ML-based constitutive modeling including the work of Sun et al. [80]. Moreover, researchers such as Koeppe et al. have developed approaches that can develop interpretable models that can be used to describe the mechanics rather than just develop statistically accurate models for constitutive response [81]. All advances reviewed here may seem like at the cutting-edge where only proof-of-concept implementations suitable for simulating very basic boundary value problems are available. While that may be true for many publications listed here, a variety of efforts have demonstrated end-to-end implementations of their approach including incorporation of the data-driven methods in FEA solvers [82, 83].

This development of data-driven model-free methods to simulate constitutive response is impressive and we believe we have all the components necessary to develop a data-driven constitutive modeling solution for Nitinol. Such an approach will allow simulation of Nitinol implant deformation based on experimental data obtained from a Nitinol base material that has undergone-specific pre-processing such pre-deformation or heat treatment steps. We propose an example framework for developing such a simulation method. The framework is schematically shown in Fig. 4. It consists of six steps.

1. The implementation begins with the creation of a standard test specimen and test protocol for acquiring experimental data used to train a data-driven constitutive law. The test specimens could be as simple as dogbone-shaped or consist of more complex forms such as cruciform or planar specimens with holes. The specimens should be suitable for acquiring spatial strain distribution history using methods such as DIC.
2. Using the standard test method described above, mechanical testing data under various conditions will be gathered. The test conditions should encompass the conditions that the modeled implant will be typically experiencing. The mechanical data could consist of local strain and global load histories.
3. The implementation then extracts the local stress field from the mechanical data recorded above. This can be achieved using the approach of Stainier et al. [72] or Cameron and Tasan [75] described above. The local stress–strain histories obtained in this step will be stored in a database and will serve as the training and validation data.
4. In this step, a data-driven model will be constructed for the constitutive response using the stress–strain history data described above and a suitable statistical method such as an ML regression scheme. The output of this

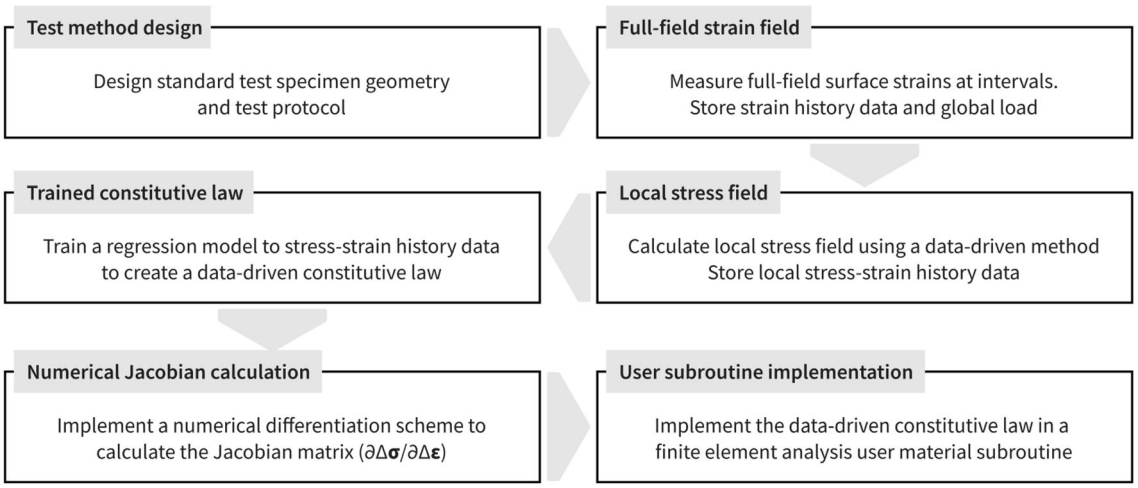

Fig. 4 A example workflow proposed for developing a data-driven constitutive law for Nitinol superelasticity

step is a database of ML regression model hyperparameters that can be stored.

5. FEA implementations of constitutive modeling typically require a constitutive law and a tangent modulus or the Jacobian matrix ($\partial\boldsymbol{\sigma}/\partial\boldsymbol{\epsilon}$) to perform the computation in a discretized implicit time-integration scheme. The Jacobian will be constructed using a suitable automatic numerical differentiation method [83, 84].
6. Once the data-driven constitutive law and the Jacobian are available, they can be programmed in an FEA framework such as the `UMAT` user material subroutine functionality in Abaqus FEA framework. Note that the data-driven constitutive law will be invoked in each increment of the simulation and furnish the stress tensor at that material point and increment based on the strain history, strain increment, and any stored data on history-dependent or internal variables.

Discussion and Summary

Data science has emerged as a multifaceted tool that can be used to advance various aspects of Nitinol implant design and simulation. We reviewed some of the recent advances in the field and provided a perspective on adapting this tool for various scenarios such as cardiovascular implant deformation characterization, durability prediction, and constitutive modeling. This paradigm shift to data-driven methods is radical and it can be facilitated by adapting a tiered approach. As an example, we propose a three-tier approach:

1. Information tier methods for acquiring data on the deformation of implants and methods for quantifying uncertainty in the data. The purpose of these methods is to collect data to be used in the approaches in the two tiers below. The emphasis on uncertainty quantification is to make sure that the credibility information becomes an integral part of any data-driven method development.
2. Knowledge tier methods to enable fusion of data from multiple sources to develop descriptive and predictive models for phenomena such as fatigue. Typically the data are combined from multiple experimental methods such as microstructural characterization, tensile testing, and fatigue testing or from experimental and simulation methods. It is common to analyze a specific aspect of a deformation phenomenon in a reductionist sense using data from a single experimental method. The purpose of methods in this tier is to combine or augment data to promote a holistic analysis of the deformation phenomenon.
3. Prediction tier methods to develop data-driven models for stress–strain response or fatigue resistance of Nitinol implants. These methods rely on data acquired and assimilated using methods in the information and knowledge tier and sequence-learning and regression methods collectively known as machine learning or deep learning. The purpose of these methods is to increase the speed and accuracy of Nitinol deformation simulation.

Many of the concepts described in these tiers may seem abstract or demanding in resources. But, there are implant-scale design and simulation activities that are already suited for adaption of these methods. One concrete example of a design activity that can benefit from this method is the probabilistic assessment of durability of medical devices. Consider a medical device that is subjected to a distribution of boundary conditions that are determined using literature data or animal, cadaver, and patient studies. A test protocol can be constructed to load the device per those boundary conditions. The cyclic strain distribution, typically used to construct a fatigue safety factor distribution, can be directly measured on the device surface using methods described in Sect. "Methods to Acquire Data and Quantify Uncertainty in the Data". Using 3D characterization methods such as computed tomography, a 3D model of the device can be constructed and the dimensional distribution can be constructed by characterizing a number of devices from various manufacturing lots. The input distribution of material parameters can be determined by testing the Nitinol raw material from multiple lots used to manufacture the device. A simulation of the device loading can be constructed using these data and the data-driven simulation methods described above. These methods will capture the nuances of the material behavior measured in the tests including intermediate phase transformation and transformation plateau hardening. The simulations can be used to augment the fatigue safety factor distribution. These data on the experimentally measured and simulated fatigue safety factor distribution and the uncertainty in them can be used in the probabilistic durability assessment of the device.

While this approach is attractive in augmenting our holistic understanding of Nitinol deformation phenomena, for advancing our predictive capabilities for Nitinol deformation simulation, and for making design exercises more efficient as described in the example above, there are conceptual and practical limitations to its adaption. These limitations must be understood so that we can set realistic expectations about the adaption of these methods and gradually overcome these limitations. The full-field experimental methods such as digital image correlation described in Sect. "Methods to Acquire Data and Quantify Uncertainty in the Data" are capital-intensive to set up and

technically demanding to use. This is particularly true for applying them to small specimens such as a typical Nitinol stent or for specimen geometries that make it difficult to get a line of sight on the area of interest such as a valve frame with a non-cylindrical shape. Improvement in optics and image data processing techniques are making the use of less sophisticated hardware more feasible, potentially alleviating the operator effort in recording the deformation data. The predictive fatigue modeling approach described in Sect. "Methods to Fuse Data and Uncover Mechanisms" can be efficiently adapted if we prioritize the gaps in our current understanding of fatigue. The effect of pre-deformation on fatigue of Nitinol raw materials is well-documented but its utility is ambiguous [85, 86]. A holistic device-level model developed using the approach described in Sect. "Methods to Fuse Data and Uncover Mechanisms" can lead to a useful guidance to the device design community. Incorporation of effects such as the fatigue of Nitinol that has undergone various cold-working sequences, or Nitinol with ternary elemental additions, or Nitinol composites such as core-shell architectures is certainly more challenging. Those challenges can be tackled as the methods to acquire and fuse data become more efficient. The practicality of data fusion depends on the speed with which the experimental or simulation input data can be gathered. The uncertainty in the fused data and predictions made using that data depend on the uncertainty in the individual data sources. The data fusion approach described here can only be practically implemented using information systems concepts such as knowledge gradient [87–89] that enable optimizing data collection to maximize the amount of information collected. These approaches to maximize information from a limited number of experiments can benefit the training of data-based constitutive methods reviewed in this work. The sequence-learning methods described in Sect. "Methods to Make Data-Driven Predictions of Nitinol Deformation Response" for developing data-driven models for the simulation of Nitinol deformation are maturing fast and readily available. Yet, the rate at which new concepts are introduced in the field, the changes in the terminology, and the changes to the available software implementations of the machine learning methods can pose challenges to the engineers implementing these data-driven models.

Overcoming these challenges and overcoming the practical limitations to the adaption of these methods will require the collective effort of the community. Consensus-building will be useful in encouraging participation in this effort. The key concepts, new outcomes, and best practices in these data-driven methods can be discussed at conferences such as the Shape Memory and Superelastic Technologies Conference (SMST), the Cardiovascular Implant Durability Conference (CVID), and the ASTM Committee Weeks to develop community-wide awareness and to promote discussion. The methods can be discussed in working groups modeled after efforts such as Best Practices for Fatigue Assessment of Heart Valve Devices organized by the Heart Valve Collaboratory and consensus best practices may be published. Some of these methods can be standardized through collaboratively developed ASTM standards. The industry participation will also be encouraged if additional research demonstrating proof-of-concept implementation of some of these methods for Nitinol becomes available. These methods inherently depend on customized software tools. Publication of such data science software in open source repositories will reduce duplication of effort across various stakeholders and reduce the overall effort required to adapt these tools. With broad collaboration between various stakeholders, we hope these data-based methods mature quickly and play a central role in the design and simulation of Nitinol cardiovascular implants.

Acknowledgments HMP would like to thank Confluent Medical Technologies, Inc. for providing the resources for this work.

References

1. Sparks TD, Kauwe SK, Parry ME, Tehrani AM, Brgoch J (2020) Machine learning for structural materials. Annu Rev Mater Res 50:27–48
2. Arróyave R, McDowell DL (2019) Systems approaches to materials design: past, present, and future. Annu Rev Mater Res 49:103–126
3. Otsuka K, Ren X (2005) Physical metallurgy of Ti-Ni-based shape memory alloys. Progress Mater Sci 50:511–678
4. Eggeler G, Hornbogen E, Yawny A, Heckmann A, Wagner M (2004) Structural and functional fatigue of NiTi shape memory alloys. Mater Sci Eng A 378:24–33
5. Robertson SW, Pelton AR, Ritchie RO (2012) Mechanical fatigue and fracture of Nitinol. Int Mater Rev 57:1–37
6. Bhattacharya K (2004) Microstructure of martensite: why it forms and how it gives rise to the shape-memory effect. Oxford series on materials modelling. Oxford University Press, Oxford
7. Paranjape HM, Bowers ML, Mills MJ, Anderson PM (2017) Mechanisms for phase transformation induced slip in shape memory alloy micro-crystals. Acta Mater 132:444–454

8. Reinholz B, Brinckmann S (2012) Phase transformations in the proximity of TiC precipitates in a NiTi matrix during fatigue. Int J Fatigue 41:72–82
9. Auricchio F, Taylor R, Lubliner J (1997) Shape-memory alloys: macromodeling and numerical simulations of superelastic behavior. Comput Methods Appl Mech Eng 146:281–312
10. Cisse C, Zaki W, Ben Zineb T (2016) A review of constitutive models and modeling techniques for shape memory alloys. Int J Plast 76:244–284
11. Zaki W, Moumni Z (2007) A 3D model of the cyclic thermomechanical behavior of shape memory alloys. J Mech Phys Solids 55:2427–2454
12. Zaki W, Zamfir S, Moumni Z (2010) An extension of the ZM model for shape memory alloys accounting for plastic deformation. Mech Mater 42:266–274
13. Wang J, Moumni Z, Zhang W, Xu Y, Zaki W (2017) A 3D finite-strain-based constitutive model for shape memory alloys accounting for thermomechanical coupling and martensite reorientation. Smart Mater Struct 26:065006
14. Peultier B, Ben Zineb T, Patoor E (2006) Macroscopic constitutive law of shape memory alloy thermomechanical behaviour. Application to structure computation by FEM. Mech Mater 38:510–524
15. ASME (2018) Assessing credibility of computational modeling through verification and validation: application to medical devices, technical report
16. Administration UFaD (2022) Assessing the credibility of computational modeling and simulation in medical device submissions. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/assessing-credibility-computational-modeling-and-simulation-medical-device-submissions. Accessed 15 Dec 2022
17. Delpueyo D, Jury A, Balandraud X, Grédiac M (2021) Applying full-field measurement techniques for the thermomechanical characterization of shape memory alloys: a review and classification. Shape Mem Superelast 7:462–490
18. Paranjape HM, Paul PP, Sharma H, Kenesei P, Park J-S, Duerig T, Brinson LC, Stebner AP (2017) Influences of granular constraints and surface effects on the heterogeneity of elastic, superelastic, and plastic responses of polycrystalline shape memory alloys. J Mech Phys Solids 102:46–66
19. Sedmák P, Pilch J, Heller L, Kopeček J, Wright J, Sedlák P, Frost M, Šittner P (2016) Grain-resolved analysis of localized deformation in nickel-titanium wire under tensile load. Science 353:559–562
20. Pan B (2018) Digital image correlation for surface deformation measurement: historical developments, recent advances and future goals. Meas Sci Technol 29:082001
21. Jones E, Iadicola M, Bigger R, Blaysat B, Boo C, Grewer M, Hu J, Jones A, Klein M, Raghavan K, Reu P, Schmidt T, Siebert T, Simenson M, Turner D, Vieira A, Weikert T (2018) International digital image correlation society. A good practices guide for digital image correlation, technical report, International Digital Image Correlation Society, ed. 1. http://idics.org/guide/. https://doi.org/10.32720/idics/gpg.ed1. Accessed 15 Dec 2022
22. Avril S, Bonnet M, Bretelle A-S, Grédiac M, Hild F, Ienny P, Latourte F, Lemosse D, Pagano S, Pagnacco E, Pierron F (2008) Overview of identification methods of mechanical parameters based on full-field measurements. Exp Mech 48:381
23. Pierron F, Grédiac M (2021) Towards material testing 2.0. A review of test design for identification of constitutive parameters from full-field measurements. Strain 57:e12370. https://doi.org/10.1111/str.12370
24. Lava P, Jones EMC, Wittevrongel L, Pierron F (2020) Validation of finite-element models using full-field experimental data: levelling finite-element analysis data through a digital image correlation engine. Strain 56:e12350. https://doi.org/10.1111/str.12350
25. Aycock KI, Weaver JD, Paranjape HM, Senthilnathan K, Bonsignore C, Craven BA (2020) Full-field microscale strain measurements of a nitinol medical device using digital image correlation. J Mech Behav Biomed Mater 114:104221
26. Senol K, Cao H, Tripathy S (2021) Characterization and validation of fatigue strains for superelastic nitinol using digital image correlation. J Med Devices 10(1115/1):4052012
27. Joung A, Pagan D, Gilbert J, Vien L, Ong I, Bonsignore C, Paranjape H (2020) A digital image correlation methodology for the characterization of cyclic deformation in nickel-titanium medical device fatigue test specimens. https://doi.org/10.13140/RG.2.2.25140.53123
28. Tripathy S, Wu M, Cao H (2019) Finite element framework for fatigue performance assessment of superelastic nitinol used in medical devices. In: Mitchell MR, Berg BT, Woods TO, Jerina KL (eds) Fourth symposium on fatigue and fracture of metallic medical materials and devices, ASTM International, 100 Barr Harbor Drive, PO Box C700, West Conshohocken, PA 19428–2959, pp 31–53. https://www.astm.org/doiLink.cgi?STP161620180039. https://doi.org/10.1520/STP161620180039
29. Cao H, Wu MH, Zhou F, McMeeking RM, Ritchie RO (2020) The influence of mean strain on the high-cycle fatigue of Nitinol with application to medical devices. J Mech Phys Solids 143:104057
30. Marrey R, Baillargeon B, Dreher ML, Weaver JD, Nagaraja S, Rebelo N, Gong X-Y (2018) Validating fatigue safety factor calculation methods for cardiovascular stents. J Biomech Eng 140:061001
31. Boukhtache S, Abdelouahab K, Berry F, Blaysat B, Grédiac M, Sur F (2021) When deep learning meets digital image correlation. Opt Lasers Eng 136:106308
32. Yang R, Li Y, Zeng D, Guo P (2022) Deep DIC: deep learning-based digital image correlation for end-to-end displacement and strain measurement. J Mater Process Technol 302:117474
33. Zhu C, Wang H, Kaufmann K, Vecchio KS (2020) A computer vision approach to study surface deformation of materials. Meas Sci Technol 31:055602
34. Freitas CJ (2020) Standards and methods for verification, validation, and uncertainty assessments in modeling and simulation. J Verif Valid Uncertain Quantif 5:021001
35. TMS (TMS) (2020) Accelerating the broad implementation of verification & validation in computational models of the mechanics of materials and structures, technical report 978-0-578-75450-5, The Materials Society
36. Ricciardi DE, Chkrebtii OA, Niezgoda SR (2019) Uncertainty quantification for parameter estimation and response prediction: generalizing the random effects Bayesian inferential framework to account for material and experimental variability, integrating materials and manufacturing. Innovation 8:273–293
37. Kennedy MC, O'Hagan A (2001) Bayesian calibration of computer models. J R Stat Soc B 63:425–464
38. Rappel H, Beex LAA, Hale JS, Noels L, Bordas SPA (2020) A tutorial on Bayesian inference to identify material parameters in solid mechanics. Arch Comput Methods Eng 27:361–385
39. Viana FAC, Subramaniyan AK (2021) A survey of Bayesian calibration and physics-informed neural networks in scientific modeling. Arch Comput Methods Eng 28:3801–3830
40. Paranjape HM, Aycock KI, Bonsignore C, Weaver JD, Craven BA, Duerig TW (2021) A probabilistic approach with built-in uncertainty quantification for the calibration of a superelastic constitutive model from full-field strain data. Comput Mater Sci 192:110357

41. Lye A, Cicirello A, Patelli E (2021) Sampling methods for solving Bayesian model updating problems: a tutorial. Mech Syst Signal Process 159:107760
42. Cranmer K, Brehmer J, Louppe G (2020) The frontier of simulation-based inference. Proc Natl Acad Sci 117:30055–30062
43. Organization IS (2021) Cardiovascular implants—cardiac valve prostheses—Part 1: general requirements (ISO 5840-1:2021). https://www.iso.org/obp/ui/#!iso:std:77033:en. Accessed 15 Dec 2022
44. Sangid MD (2020) Coupling in situ experiments and modeling—opportunities for data fusion, machine learning, and discovery of emergent behavior. Curr Opin Solid State Mater Sci 24:100797
45. Chen J, Liu Y (2022) Fatigue modeling using neural networks: a comprehensive review. Fatigue Fract Eng Mater Struct 45:945–979. https://doi.org/10.1111/ffe.13640
46. Durmaz AR, Hadzic N, Straub T, Eberl C, Gumbsch P (2021) Efficient experimental and data-centered workflow for microstructure-based fatigue data. Exp Mech 61:1489–1502
47. Gebhardt C, Trimborn T, Weber F, Bezold A, Broeckmann C, Herty M (2020) Simplified ResNet approach for data driven prediction of microstructure-fatigue relationship. Mech Mater 151:103625
48. Murakami Y, Endo M (1994) Effects of defects, inclusions and inhomogeneities on fatigue strength. Int J Fatigue 16:163–182
49. Murakami Y (2019) Metal fatigue—effects of small defects and nonmetallic inclusions, 2nd edn, Academic Press. https://linkinghub.elsevier.com/retrieve/pii/B9780128138762000406. https://doi.org/10.1016/B978-0-12-813876-2.00040-6
50. Murakami Y (1989) Quantitative evaluation of effects of nonmetallic inclusions on fatigue strength of high strength steels. I: Basic fatigue mechanism and evaluation of correlation between the fatigue fracture stress and the size and location of nonmetallic inclusions. Int J Fatigue 11:8
51. Murakami Y, Usuki H (1989) Quantitative evaluation of effects of non-metallic inclusions on fatigue strength of high strength steels. II: Fatigue limit evaluation based on statistics for extreme values of inclusion size. Int J Fatigue 11:299–307
52. Beretta S (2021) More than 25 years of extreme value statistics for defects: fundamentals, historical developments, recent applications. Int J Fatigue 151:106407
53. Urbano MF, Cadelli A, Sczerzenie F, Luccarelli P, Beretta S, Coda A (2015) Inclusions size-based fatigue life prediction model of NiTi alloy for biomedical applications. Shape Mem Superelast 1:240–251
54. Kafka OL, Yu C, Shakoor M, Liu Z, Wagner GJ, Liu WK (2018) Data-driven mechanistic modeling of influence of microstructure on high-cycle fatigue life of nickel titanium. JOM 70:1154–1158
55. Moore JA, Frankel D, Prasannavenkatesan R, Domel AG, Olson GB, Liu WK (2016) A crystal plasticity-based study of the relationship between microstructure and ultra-high-cycle fatigue life in nickel titanium alloys. Int J Fatigue 91:183–194
56. Priddy MW, Paulson NH, Kalidindi SR, McDowell DL (2017) Strategies for rapid parametric assessment of microstructure-sensitive fatigue for HCP polycrystals. Int J Fatigue 104:231–242
57. Paulson NH, Priddy MW, McDowell DL, Kalidindi SR (2018) Data-driven reduced-order models for rank-ordering the high cycle fatigue performance of polycrystalline microstructures. Mater Des 154:170–183
58. Barbosa JF, Correia JAFO, Júnior RCSF, Jesus AMPD (2020) Fatigue life prediction of metallic materials considering mean stress effects by means of an artificial neural network. Int J Fatigue 135:105527
59. Chen J, Liu Y (2021) Probabilistic physics-guided machine learning for fatigue data analysis. Expert Syst Appl 168:114316
60. Dourado A, Irmak F, Viana FAC, Gordon AP (2021) A nonstationary uncertainty model and Bayesian calibration of strain-life models. J Verif Valid Uncertain Quantif 10(1115/1):4049324
61. Kovachki N, Liu B, Sun X, Zhou H, Bhattacharya K, Ortiz M, Stuart A (2022) Multiscale modeling of materials: computing, data science, uncertainty and goal-oriented optimization. Mech Mater 165:104156
62. Ghaboussi J, Garrett JH, Wu X (1991) Knowledge-based modeling of material behavior with neural networks. J Eng Mech 117:132–153
63. Mozaffar M, Bostanabad R, Chen W, Ehmann K, Cao J, Bessa MA (2019) Deep learning predicts path-dependent plasticity. Proc Natl Acad Sci 116:26414–26420
64. Gorji MB, Mozaffar M, Heidenreich JN, Cao J, Mohr D (2020) On the potential of recurrent neural networks for modeling path dependent plasticity. J Mech Phys Solids 143:103972
65. Abueidda DW, Koric S, Sobh NA, Sehitoglu H (2021) Deep learning for plasticity and thermo-viscoplasticity. Int J Plast 136:102852
66. Bonatti C, Mohr D (2021) On the importance of self-consistency in recurrent neural network models representing elasto-plastic solids. J Mech Phys Solids 158:104697
67. Ciftci K, Hackl K (2022) Model-free data-driven simulation of inelastic materials using structured data sets, tangent space information and transition rules. Comput Mech 70:425–435
68. Huang D, Fuhg JN, Weißenfels C, Wriggers P (2020) A machine learning based plasticity model using proper orthogonal decomposition. Comput Methods Appl Mech Eng 365:113008
69. Liu B, Kovachki N, Li Z, Azizzadenesheli K, Anandkumar A, Stuart AM, Bhattacharya K (2022) A learning-based multiscale method and its application to inelastic impact problems. J Mech Phys Solids 158:104668
70. Karapiperis K, Stainier L, Ortiz M, Andrade JE (2020) Data-driven multiscale modeling in mechanics. J Mech Phys Solids 147:104239
71. Eggersmann R, Kirchdoerfer T, Reese S, Stainier L, Ortiz M (2019) Model-free data-driven inelasticity. Comput Methods Appl Mech Eng 350:81–99
72. Stainier L, Leygue A, Ortiz M (2019) Model-free data-driven methods in mechanics: material data identification and solvers. Comput Mech 64:381–393
73. Réthoré J, Leygue A, Coret M, Stainier L, Verron E (2018) Computational measurements of stress fields from digital images: computational measurements of stress fields from digital images. Int J Numer Methods Eng 113:1810–1826
74. Langlois R, Coret M, Réthoré J (2022) Non-parametric stress field estimation for history-dependent materials: application to ductile material exhibiting Piobert-Lüders localization bands. Strain 58:e12410
75. Cameron BC, Tasan CC (2021) Deterministic calculation of elasto-plastic stress-strain behavior from arbitrary deformation fields. arXiv:2103.11938 [cond-mat]
76. Flaschel M, Kumar S, De Lorenzis L (2022) Discovering plasticity models without stress data. arXiv preprint arXiv:2202.04916
77. Flaschel M, Kumar S, De Lorenzis L (2021) Unsupervised discovery of interpretable hyperelastic constitutive laws. Comput Methods Appl Mech Eng 381:113852
78. Masi F, Stefanou I, Vannucci P, Maffi-Berthier V (2021) Thermodynamics-based artificial neural networks for constitutive modeling. J Mech Phys Solids 147:104277
79. Karniadakis GE, Kevrekidis IG, Lu L, Perdikaris P, Wang S, Yang L (2021) Physics-informed machine learning. Nat Rev Phys 3:422–440

80. Sun X, Bahmani B, Vlassis NN, Sun W, Xu Y (2021) Data-driven discovery of interpretable causal relations for deep learning material laws with uncertainty propagation. Granul Matter 24:1
81. Koeppe A, Bamer F, Selzer M, Nestler B, Markert B (2022) Explainable artificial intelligence for mechanics: physics-explaining neural networks for constitutive models. Front Mater 8:824958
82. Liu D, Yang H, Elkhodary KI, Tang S, Liu WK, Guo X (2022) Mechanistically informed data-driven modeling of cyclic plasticity via artificial neural networks. Comput Methods Appl Mech Eng 393:114766
83. Huang DZ, Xu K, Farhat C, Darve E (2020) Learning constitutive relations from indirect observations using deep neural networks. J Comput Phys 416:109491
84. Xu K, Darve E (2020) ADCME: learning spatially-varying physical fields using deep neural networks. http://arxiv.org/abs/2011.11955. https://doi.org/10.48550/arXiv.2011.11955, arXiv:2011.11955 [cs, math]
85. Gupta S, Pelton AR, Weaver JD, Gong X-Y, Nagaraja S (2015) High compressive pre-strains reduce the bending fatigue life of nitinol wire. J Mech Behav Biomed Mater 44:96–108
86. Senthilnathan K, Shamimi A, Bonsignore C, Paranjape H, Duerig T (2019) Effect of prestrain on the fatigue life of superelastic nitinol. J Mater Eng Perform 28:5946–5958
87. Ghoreishi SF, Allaire DL (2018) A fusion-based multi-information source optimization approach using knowledge gradient policies. In: 2018 AIAA/ASCE/AHS/ASC structures, structural dynamics, and materials conference, American Institute of Aeronautics and Astronautics, Kissimmee, Florida. https://doi.org/10.2514/6.2018-1159
88. Honarmandi P, Duong TC, Ghoreishi SF, Allaire D, Arroyave R (2019) Bayesian uncertainty quantification and information fusion in CALPHAD-based thermodynamic modeling. Acta Mater 164:636–647
89. Ghoreishi SF, Molkeri A, Arróyave R, Allaire D, Srivastava A (2019) Efficient use of multiple information sources in material design. Acta Mater 180:260–271