

Mechanical fatigue and fracture of Nitinol

S. W. Robertson*¹, A. R. Pelton¹ and R. O. Ritchie²

Nitinol, a near equiatomic intermetallic of nickel and titanium, is the most widely known and used shape memory alloy. Owing to its capacity to undergo a thermal or stress induced martensitic phase transformation, Nitinol displays recoverable strains that are more than an order of magnitude greater than in traditional alloys, specifically as high as 10%. Since its discovery in the 1960s, Nitinol has been used for its shape memory properties for couplings and actuators, although its contemporary use has been in for medical devices. For these applications, the stress induced transformation ('superelasticity') has been used extensively for self-expanding implantable devices such as endovascular stents and *vena cava* filters, and for tools such as endodontic files. Most of these applications involve cyclically varying biomechanical stresses or strains that drive the need to fully understand the fatigue and fracture resistance of this alloy. Here we review the existing knowledge base on the fatigue of Nitinol, both in terms of their stress or strain life (total life) and damage tolerant (crack propagation) behaviour, together with their fracture toughness properties. We further discuss the application of such data to the fatigue design and life prediction methodologies for Nitinol implant devices used in the medical industry.

Keywords: Nitinol, Fatigue, Fracture, Toughness, Total life, Damage tolerant, Review

Glossary of terms

Material specific terms

Austenite – stable phase of Nitinol at temperatures above A_f with a simple cubic B2 (CsCl) crystal structure.

Austenite finish temperature A_f – the temperature above which Nitinol is fully austenitic.

Austenite start temperature A_s – the temperature at which, upon heating, martensitic Nitinol begins to phase transform to austenite.

Martensite – stable phase of Nitinol at temperatures below M_f with a monoclinic B19' crystal structure. This phase can be stress induced from austenite at temperatures between A_s and M_d .

Martensitic deformation temperature M_d – the temperature at which the energy to stress induce martensite is equal to the energy to deform austenite plastically. Thus, the M_d coincides with temperature at which the stress induced phase transformation is suppressed.

Martensite finish temperature M_f – the temperature below which Nitinol is fully martensitic.

Martensite start temperature M_s – the temperature at which, upon cooling, austenitic Nitinol begins to phase transform to martensite.

Nitinol Ni_XTi_Y – near equiatomic intermetallic of nickel (X at-%) and titanium (Y at-%) capable of exhibiting superelastic behaviour due to a stress induced

phase transformation, or shape memory due to a temperature induced phase transformation.

R phase – a rhombohedral phase of Nitinol that forms from the cubic austenite phase and is considered another martensite with a $P\bar{3}$ crystal structure. The R phase often competes with the austenite–B19' martensite transformation.

Shape memory – ability of a material to return to a pretrained geometry upon the application of heat. This is accomplished by cooling the material below M_s , then deforming it by detwinning of martensite, followed by raising the temperature to above A_f at which point it returns to its original undeformed geometry. Shape memory Nitinol can recover up to 10% strain.

Stress induced martensite – B19' martensite formed from the parent austenite phase in Nitinol through the application of stress energy into the material at temperatures above A_s . It is also possible to stress induce $P\bar{3}$ R phase.

Superelasticity (pseudoelasticity) – ability of a material to accommodate global strains due to strain accommodation via a martensitic phase transformation; see 'stress induced martensite'. For example, superelastic Nitinol alloys can exhibit fully reversible strains up to 10%.

Thermal martensite – martensite formed from the parent austenite phase through the application of thermal energy (cooling) into the material.

Test method specific terms

Strain amplitude ϵ_a – absolute value of the difference between the mean strain and either the maximum or minimum strain applied in cyclic fatigue conditions. $\epsilon_a = |\epsilon_m - \epsilon_{\min}| = |\epsilon_m - \epsilon_{\max}|$; alternatively, the strain amplitude is half of the strain range $= |\epsilon_{\max} - \epsilon_{\min}|/2$. Since

¹Nitinol Devices and Components (NDC), 47533 Westinghouse Dr., Fremont, CA 94539, USA

²Materials Sciences Division, Lawrence Berkeley National Laboratory, and Department of Materials Science and Engineering, University of California, Berkeley, CA 94720, USA

*Corresponding author, email scott.robertson@nitinol.com

strains are second order tensors, these strain conditions apply only to conditions in which there is only one dominant cyclic tensile strain or tensile stress and it is in a constant material direction.

Stress amplitude σ_a – absolute value of the difference between the mean stress and either the maximum or minimum stress applied in cyclic fatigue conditions. $\sigma_a = |\sigma_m - \sigma_{\min}| = |\sigma_m - \sigma_{\max}|$; alternatively, the stress amplitude is half of the stress range $= |\sigma_{\max} - \sigma_{\min}|/2$. Since stresses are second order tensors, these stress conditions apply only to conditions in which there is only one dominant cyclic tensile strain or tensile stress and it is in a constant material direction.

Critical crack length a_{crit} – dimensional length of a flaw below which the material will not preferentially propagate a fatigue crack from that flaw. Flaws longer than a_{crit} serve as stress concentrators in a material and act as preferential nucleation sites for fatigue crack propagation.

Damage tolerant fatigue – measurement of fatigue crack growth from a flaw or precrack that is intentionally introduced into the specimen. This approach measures only the propagation stage of the fatigue life.

Displacement control – total life cyclic fatigue test run by fixing the displacement amplitudes regardless of force required to achieve those displacements; a.k.a. strain control.

Fracture toughness K_C – engineering parameter used to describe the critical stress intensity factor that causes fracture when a monotonic load is applied to the material. The fracture toughness is often dependent upon the length of the flaw from which fracture originates, and usually asymptotes to a stable toughness value above some flaw length. This dependence of fracture toughness on the flaw size is characterised by a crack resistance curve (R curve).

Load control – total life cyclic fatigue test run by fixing the alternating loads regardless of displacement required to achieve those loads; a.k.a. stress control.

Load ratio (R ratio) – ratio of the minimum to maximum loads (P_{\min}/P_{\max}), and hence for tensile loading, it is the ratio of the minimum to maximum stress intensities (K_{\min}/K_{\max}). See also ‘strain ratio (R_ϵ ratio)’ below.

Mean strain ϵ_m – average strain value between the maximum and minimum applied strain during cyclic loading conditions; $\epsilon_m = (\epsilon_{\max} + \epsilon_{\min})/2$.

Mean stress σ_m – average stress value between the maximum and minimum applied stress during cyclic loading conditions; $\sigma_m = (\sigma_{\max} + \sigma_{\min})/2$.

Paris law – equation used to describe the linear portion of a fatigue crack growth curve per $da/dN = C\Delta K^m$, where a is the instantaneous flaw length, N is the number of fatigue cycles, C and m are scaling constants which are dependent upon the material in question and the loading/environmental conditions, e.g. the R ratio, the cyclic frequency, the external environment, etc.

Plane strain – a boundary condition such that one of the three principal strain components applied to the specimen is zero. This condition exists when one dimension is significantly large enough to effectively constrain the material within that axis, thereby preventing strain in that dimension.

Plane stress – a boundary condition such that one of the three principal stress components applied to the specimen is zero. This condition exists in specimens with

significantly small dimensions in one axis. Moreover, the surfaces of specimens may be considered plane stress since the material is free to deform into open space without any constraint.

R curve – the crack resistance curve, describing the resistance to fracture plotted as a function of subcritical crack extension Δa .

Strain ratio R_ϵ – ratio of the minimum to maximum strains ($\epsilon_{\min}/\epsilon_{\max}$), used similarly to the load ratio R . In the case of non- $R = -1$ conditions, it is appropriate to cite the mean strain and strain amplitude for clarification.

Stress intensity factor K – linear elastic field parameter used to describe the stress and displacement fields in the immediate vicinity of a crack-like flaw in an elastic material. K is proportional to the applied stress σ_{app} , and square root of the flaw size a as per the following expression: $K = Q\sigma_{\text{app}}(\pi a)^{1/2}$, where Q is a geometry factor of order unity.

Stress intensity range ΔK – the difference between the maximum and minimum stress intensity factors $K_{\max} - K_{\min}$ that a specimen witnesses during a given fatigue loading cycle.

Stress intensity threshold ΔK_{th} – the stress intensity range value below which fatigue crack propagation occurs at vanishingly small growth rates (typically below $\sim 10^{-10}$ to 10^{-11} m/cycle).

Stress range $\Delta\sigma$ – difference between the maximum and minimum applied stresses. $\Delta\sigma = (\sigma_{\max} - \sigma_{\min}) = 2\sigma_a$.

Strain range $\Delta\epsilon$ – difference between the maximum and minimum applied strains. $\Delta\epsilon = (\epsilon_{\max} - \epsilon_{\min}) = 2\epsilon_a$.

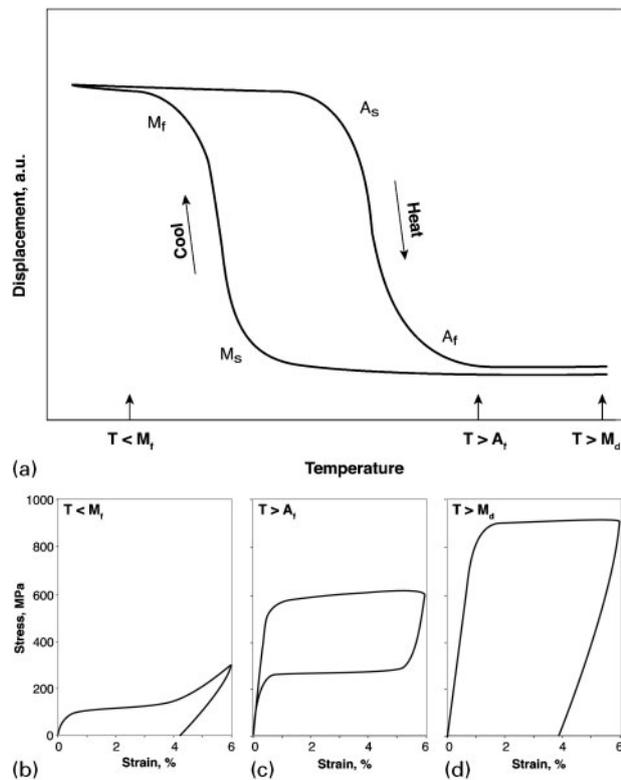
Total life fatigue – measurement of fatigue crack growth from a specimen that has undergone treatment (e.g. polishing and heat treatment) to render the surface comparable with a commercially equivalent device without intentional introduction of a flaw or precrack. This approach measures both the nucleation and propagation stages of the fatigue life, where nucleation originates from stress concentrations native to the material (e.g. grain boundary and inclusion).

Introduction

‘Nitinol’ (NiTi) is a nearly equiatomic intermetallic of nickel and titanium. Its unique properties were first revealed in 1962 by Buehler and Wang at the Naval Ordnance Laboratory – indeed, the term ‘Nitinol’ is the acronym for Nickel Titanium Naval Ordnance Laboratory. The commercial importance of Nitinol is due to its closely related properties of shape memory and superelasticity (also referred to as pseudoelasticity). Both properties result from the phase transformation between an austenitic (parent) phase [with a simple cubic B2 (CsCl) structure] to a martensite (daughter) phase (with a monoclinic B19’ structure)* as described in several references.²⁻⁴ Although some commercial success has been derived from the shape memory applications (e.g. aircraft fuel line couplings and actuator springs), superelastic applications have garnered much greater success, especially for medical devices.^{5,6} Therefore, this manuscript focuses mostly upon the superelastic rather than shape memory phase transformations as they pertain to the fatigue and fracture resistance of Nitinol.

The mechanical properties of Nitinol depend dramatically on the deformation temperatures. For example,

*Another martensitic phase, known as the R phase, has a rhombohedral crystal structure and also shows shape memory and superelasticity.^{1,2}



1 a schematic displacement–temperature curve under constant tensile force for Nitinol that indicates the transformation temperatures as well as the three general regions for which fatigue data have been generated. The monotonic stress–strain curves at test temperatures: *b* below M_f (100% martensite), which is characterised with a low stress plateau and a large remnant strain after unloading from 6% strain; *c* above A_f (100% austenite), which is characterised by an increased upper stress plateau (transformation from austenite to deformed martensite), an unloading plateau (transformation from martensite to austenite) and a return to the original shape; *d* above M_d , the temperature above which plasticity is the dominant deformation accommodation mechanism. After Pelton *et al.*⁷

Fig. 1 shows the idealised relationship between the constant force strain–temperature curve that illustrates the shape memory effect and the uniaxial tensile stress–strain curves that result at the three specified deformation temperatures.⁷ As shown in Fig. 1a, martensite is formed upon cooling from the austenite phase with a concomitant increase in strain. The temperature at which this transformation begins is referred to as the martensite start temperature M_s and the temperature at which the transformation ends is referred to as the martensite finish temperature M_f . The strain begins to recover upon heating this deformed martensite above the austenite start temperature A_s with recovery of the strain and the austenite phase at the austenite finish temperature A_f . At intermediate temperatures between M_s and M_f or A_s and A_f , there is a two phase mixture of austenite and martensite. There are a total of 24 variants of martensite that develop by cooling from austenite.^{2,4} Figure 1b shows the uniaxial stress–strain curve of the fully martensite phase (below M_f). Here the initially twinned martensite deforms to produce detwinned and deformed

martensite. Note that the 6% strain is not fully recovered when the stress is removed; full recovery is only expected upon heating above A_f . Figure 1c illustrates the phenomena of stress induced martensite (superelasticity) expected at temperatures above A_f . The majority of medical devices are superelastic at body temperature, so this condition is described more fully below.

Superelasticity involves the stress induced transformation, followed by strain recovery (and concomitant phase transformation back to the parent austenite) upon removal of the applied stress. First, the initial linear–elastic deformation is characterised by a monotonic increase in stress over a small strain range up to the critical transformation stress. After $\sim 1\%$ strain, the curve flattens to form a plateau with only a small change in stress over a much larger strain range during which the austenite transforms to detwinned (deformed) martensite. The martensite laths nucleate and grow with the preferred martensite variant according to the relevant Taylor factor (inverse Schmid factor) that relieves the internal stress.^{2,4} Increasing deformation in this range increases the volume fraction of stress induced martensite, whereby at the end of the stress plateau, the stress induced martensitic transformation is (nominally) complete. Recovery of up to 10% strains is possible, which is a full order of magnitude greater than exhibited by, for example, stainless steel, with its elastic strain limit of $\sim 0.2\%$. Upon unloading the applied stress, the martensite becomes unstable and the reverse transformation occurs with the stress stabilised martensite reverting along the previous crystallographic path to austenite; there is a corresponding decrease in martensite volume fraction and a return to its original configuration. Although not shown in Fig. 1c, it is important to note that beyond the loading plateau, deformation of the stress induced martensite continues whereupon elastic followed by plastic (dislocation) yielding occurs. In contrast to deformation of thermal martensite (Fig. 1b) and superelastic austenite (Fig. 1c), deformation at temperatures above M_d is accommodated by traditional plasticity and there are no martensitic transformations involved. The corresponding uniaxial stress–strain curve for these conditions is shown in Fig. 1d, whereby there is an initial linear–elastic region followed by traditional plastically above $\sim 1\%$ strain. The cyclic deformation of Nitinol is decidedly more complex than shown for these three monotonic conditions as will be discussed throughout the remainder of this paper.

Lifetime prediction of components that are subjected to cyclic mechanical motion is critical for the design and optimisation of all engineering devices manufactured from Nitinol. Medical devices, many of which are permanently implanted and experience millions to billions of *in vivo* cycles, dominate the current Nitinol market. The safety and durability of these devices may be measured by their fatigue and fracture resistance. Descriptions of mechanical fatigue on a microscopic, and even macroscopic, level for these Nitinol based medical devices are complicated and remain incomplete even today. The source of the complications is due to the uncertain role of the transformation under cyclically varying deformations and the complexity of the various phases in Nitinol (including austenite, martensite, R phase, oxide and carbide inclusions, as well as Ni rich and Ti rich precipitates).²

During the past 40 years, there have been many investigations of both the thermal and mechanical fatigue behaviour of Nitinol. Nevertheless, taken as a whole these studies have failed to provide a comprehensive description and reliable constitutive relationships that may be universally applied for a variety of applications. In particular, an accurate description of the mechanical properties under multiaxial loading conditions is still uncertain due to the path dependence of the phase transformation. This dearth of understanding, however, is perhaps not unexpected since the ability to apply conventional stress analysis, fracture mechanics and fatigue design principles to the complex mechanical behaviour of Nitinol is not straightforward. Furthermore, this situation is compounded by the fact that much of the current medical device related fatigue investigations of Nitinol devices are invariably tied into proprietary, unpublished submissions to the Food and Drug Administration or other notified bodies. Consequently, a comprehensive approach to the fatigue design of Nitinol components, with appropriate test methods, life prediction strategies and mechanistic understanding, is simply not yet available.

Studies on the fatigue properties of Nitinol date back to the 1960s. Such early investigations, however, suffered from several problems. First, Nitinol is difficult to melt and requires tight controls on absolute purity and the reproducibility of melt compositions. These early alloys were often substandard compared with modern practices; therefore, the results of these early studies have to be used with caution if they are included to provide predictions about device fatigue. Second, many of the early investigations focused on materials that were tested in the fully annealed condition or with high temperature, long duration thermal treatments, reflecting the state-of-the-art in the 1960s through the mid-1980s. Again, the results from these studies may lead to generalisations that may not pertain to more modern thermomechanically processed materials used in Nitinol implants. Third, there is the question of how the fatigue properties were evaluated. There are three basic (and somewhat complementary) approaches that are used to evaluate the fatigue behaviour of Nitinol. These methods attempt to predict relationships among various factors (e.g. applied stress, strain, temperature and environment), or more pragmatically, solve an *ad hoc* problem. The traditional means of assessing fatigue resistance is to monitor the total lifetime of a specimen as a function of the applied stresses; this is referred to as the stress–life or S – N approach and dates back to Wöhler in the nineteenth century.⁸ A second common, and perhaps more appropriate, approach for fatigue analysis of fine featured Nitinol medical devices, is to control the deformation (strain) per cycle; this is referred to as the total strain or strain–life (ϵ – N) method and was developed by Coffin and Tavernelli,⁹ Manson¹⁰ and subsequently Morrow,¹¹ starting in the 1950s. Both the stress–life and strain–life methodologies are based on the assessment of total lifetimes, although in many instances, e.g. in the presence of defects, it can be important to distinguish between the number of cycles spent in crack initiation versus crack growth. To address this issue, a third, quite distinct, approach to fatigue design and assessment has been

used based on fracture mechanics. This ‘damage tolerant’ approach originated with Paris and co-workers in the 1960s,¹² and assumes a certain distribution of preexisting cracks and considers the growth of a single dominant crack as a function of number of cycles and applied stress. The resulting data are usually presented as log–log plots of the growth rate per cycle, da/dN , as a function of the stress intensity range ΔK computed from the stress range $\Delta\sigma$, crack length a and geometry. However, given the divergence of these various fatigue testing approaches, it is often difficult to extrapolate from one test condition to the next, especially with widely varying test specimen geometries and processing histories.

The intent of this paper is to form a more universal understanding of the inherent fatigue and fracture resistance of Nitinol particularly for its application to medical implant devices. We present here a comprehensive review of the literature on the mechanical cycling of binary Nitinol based on both total life and damage tolerant approaches. We further provide an evaluation of the corresponding fracture resistance as characterised in terms of their fracture toughness behaviour. In our assessment of research on the fatigue and fracture of Nitinol that has been reported to date, we have noted where we believe that there are obvious discrepancies in the data or other reasons for caution. In several instances we have computed fatigue and fracture design parameters that were not calculated by the original authors. Moreover, in the tabulation of these data, we have tried to document details on the transformation temperatures, test temperatures and chemical composition, when available, as these factors are now known to critically affect the mechanical, and therefore cyclic, behaviour of Nitinol. Finally, we provide a discussion of the use of these data, and the specific fatigue analysis strategies of stress–life, strain–life and damage tolerance, to the fatigue design and life prediction of biomedical Nitinol devices.

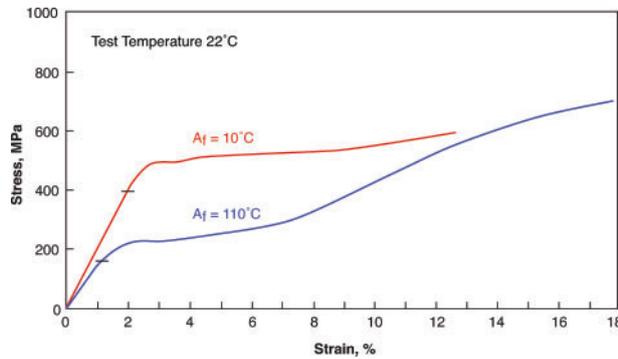
Total life fatigue

This section reviews the existing total life literature and provides insights into the effects of S – N versus ϵ – N methodologies as well as effects of test temperature, transformation temperature and applied mean stress or strain. Since total life tests are experimentally easy to conduct, the literature is rich with examples, albeit often with different starting materials, different test conditions and therefore different outcomes. It should be noted that the defining experimental study to tie these critical effects together has not been yet conducted. Consequently, when possible, we will emphasise the known factors from each study as well as speculate on the mechanisms based on the multitude of results.

Stress–life uniaxial fatigue

Zero mean stress conditions ($R = -1$)

Melton and Mercier^{13,14} (hereafter referred to as MM1 and MM2 respectively) were early pioneers in the field of the mechanical fatigue of Nitinol; no investigations before or since have attempted to address all three modes of fatigue behaviour represented by the S – N , ϵ – N and damage tolerant approaches. As part of their studies, MM1 cast binary compositions with A_f values

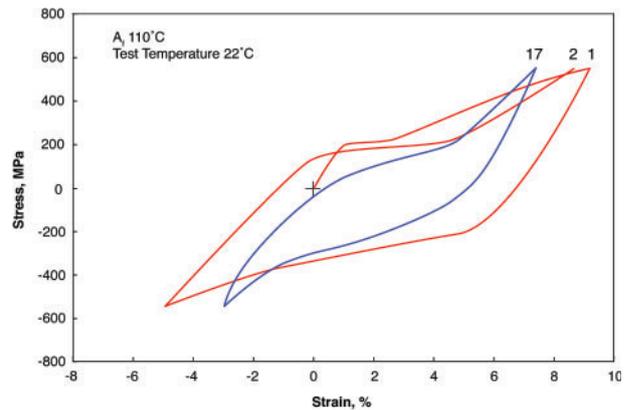


2 Uniaxial tensile stress–strain curves at room temperature for nominal compositions of $\text{Ni}_{49.9}\text{Ti}_{50.1}$ ($A_f = +10^\circ\text{C}$) superelastic austenite and $\text{Ni}_{49.7}\text{Ti}_{50.3}$ ($A_f = +110^\circ\text{C}$) thermal martensite after full anneal conditions. The 10^7 cycles fatigue limits from Fig. 5 are shown with the horizontal lines on the linear elastic regions on the respective curves. After MM1¹³

of $+10^\circ\text{C}$ ($M_s = -30^\circ\text{C}$) to $+110^\circ\text{C}$ ($M_s = 70^\circ\text{C}$), based on stress free electrical resistivity measurements[†]. All mechanical tests were conducted at room temperature, so that conditions of both austenite and thermal martensite could be tested within the range of compositions. Figure 2 shows the monotonic tensile stress–strain curves for the two extremes of their compositions, specifically (in at-%) $\text{Ni}_{49.9}\text{Ti}_{50.1}$ ($A_f = +10^\circ\text{C}$) and $\text{Ni}_{49.7}\text{Ti}_{50.3}$ ($A_f = +110^\circ\text{C}$) that were respectively superelastic austenite and thermal martensite when tested at room temperature. Note that, as expected, the superelastic composition has a higher stress plateau (corresponding to stress induced martensite formation) than the thermal martensite structure (corresponding to martensite detwinning and deformation);^{2,3} the elastic modulus of the austenite is also greater than that of the martensite. Tensile strengths for the two conditions were low (<700 MPa), due to a 1 h anneal at 950°C and the very low degree of prior deformation (17.5 mm as cast diameter to 8 mm hot forged diameter resulting in $\sim 80\%$ total reduction in area). Additionally, the tensile fracture strain (used here as a relative measure since the more accurate reduction in area measurements were not provided) was extremely low ($<20\%$) for both annealed alloys. In contrast, tensile fracture strains of $>50\%$ are common for modern annealed Nitinol.¹⁵ It is therefore possible that the MM alloys contained a high impurity (oxygen and/or carbon) content or perhaps other inhomogeneities in the melts. Furthermore, the extent of prior diametral deformation was insufficient to refine the as cast microstructure. The importance of processing and alloy impurities on mechanical and fatigue properties will be discussed in the following sections.

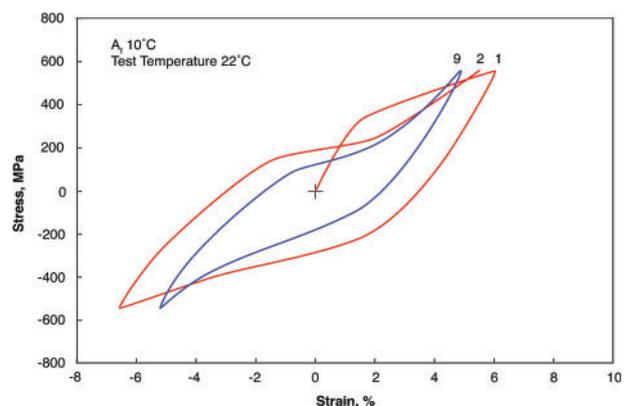
MM1 illustrated the change in stress–strain behaviour with stress controlled low cycle fatigue studies. Figure 3 shows examples of cyclic stress–strain curves of thermal martensite ($A_f = +110^\circ\text{C}$) composition that were obtained under constant stress cycling between a tensile stress of 550 MPa and a compressive stress of -550 MPa (well

[†] Through the 1980s, it was convention to characterise Nitinol with the martensite rather than austenite transformation temperatures. However, A_f has been universally used since then. Consequently, in the absence of a reported value, we approximate A_f as $M_s + 40^\circ\text{C}$. Furthermore, there are a variety of methods used to characterise the transformation temperatures, each of which yields different values. Therefore, we will cite the method when available.

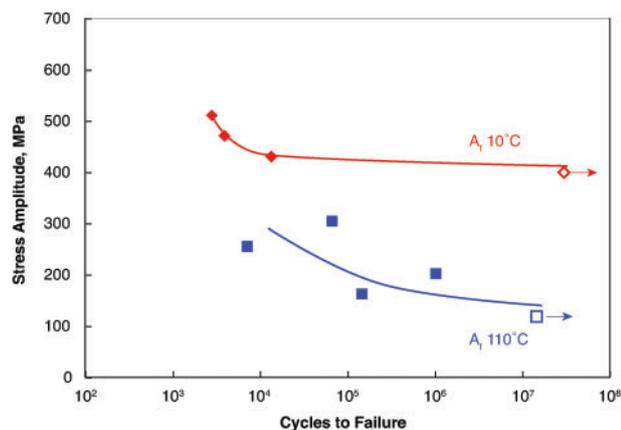


3 Cyclic stress–strain curves for Nitinol obtained at room temperature under constant stress (± 550 MPa) tension–compression conditions for annealed thermal martensite ($A_f = +110^\circ\text{C}$) composition. The numbers above the curves refer to the cycle numbers. Note that both the hysteresis and strain decrease with increasing cycles. After MM1¹³

beyond the detwinning stress plateau of $\sim 2\%$ strain in the first cycle). The curves illustrate that more strain is obtained on the first tension part of the cycle than in the corresponding compression part, so that after several cycles the hysteresis loop at zero stress is offset in the tensile direction, i.e. the specimen elongates with a resultant permanent shape change. Note also that the hysteresis loop after 17 cycles is significantly narrower than in the initial cycle. A similar result was obtained for cyclic tests on superelastic austenite with $A_f = 10^\circ\text{C}$, except that measured strains were lower (Fig. 4). When compared with the monotonic tensile stress–strain behaviour (Fig. 2), a stress amplitude of 550 MPa is still on the stress induced martensite plateau, where one would presume that only elastic and stress induced (superelastic) mechanisms are operating. However, the observation of a dynamic hysteresis curve with increasing fatigue cycles indicates that microstructural changes have occurred in both thermal martensite cycling as well as superelastic austenite cycling.



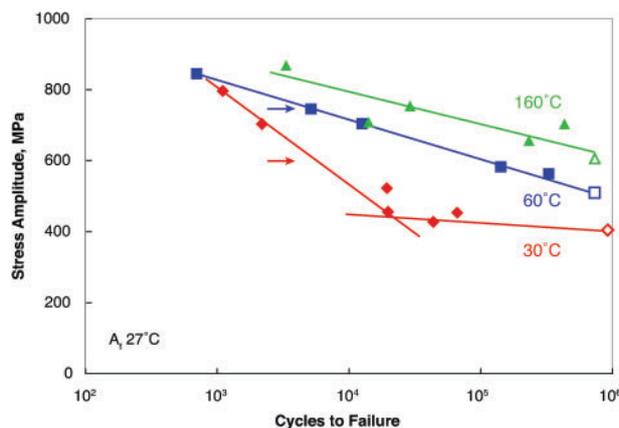
4 Cyclic stress–strain curves obtained at room temperature under constant stress (± 550 MPa) tension–compression conditions for superelastic austenite ($A_f = 10^\circ\text{C}$). The numbers above the curves refer to the cycle numbers. Note that both the hysteresis and strain decrease with increasing cycles. After MM1¹³



5 *S-N* curves for Nitinol showing the stress amplitude as a function of the number of cycles to failure. Fully reversed ($R=-1$) stress based fatigue curves out to $\sim 10^7$ cycles are shown for nominal compositions of $\text{Ni}_{49.9}\text{Ti}_{50.1}$ ($A_f=10^\circ\text{C}$) and $\text{Ni}_{49.7}\text{Ti}_{50.2}$ ($A_f=110^\circ\text{C}$), i.e. for superelastic and thermal martensite respectively. Note that the fatigue limit for the austenite is significantly higher than the martensite, commensurate with the values of their respective plateau stresses (Fig. 2). Runout conditions are noted with open symbols. After MM1¹³

Although microstructural characterisation was not part of the Melton and Mercier investigations, other studies provide some insight in the effects of fatigue damage accumulation. For example, optical metallographic images of Nitinol after progressively increasing numbers of mechanical cycles showed stress induced martensite forming during the loading cycles, with evidence of residual martensite after unloading, after as few as 10 fatigue cycles.¹⁶⁻¹⁸ Localised plastic deformation was observed in the vicinity of the martensitic plates that form, concomitant with increasing permanent deformation after each cycle.¹⁹ The martensite variants activated in each grain were consistent with each cycle and varied only in exact spatial location with subsequent cycles. However, due to the increased localised deformation, additional variants formed at each cycle, which may explain the cyclic strain hardening observed macroscopically in the MM studies. It was also demonstrated that although residual martensite plates likely alter the local stress fields that facilitate transformation, the first martensitic plates did not appear at the exact location of the residual martensite.¹⁸ The influence of transformation induced plasticity and microstructural rearrangements will be discussed in the context of strain controlled tests in the section on 'Non-zero mean strain conditions'.

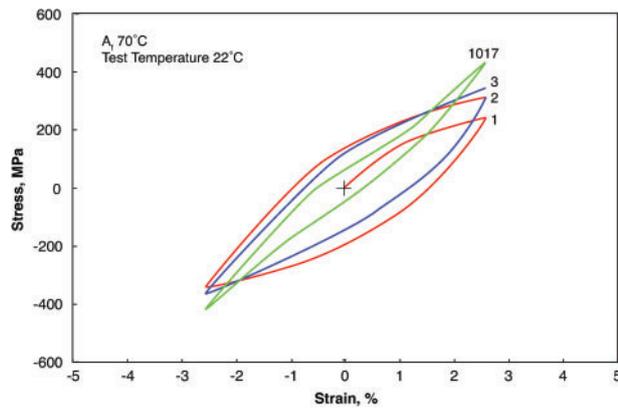
S-N curves (Fig. 5) for alloys from Figs. 2-4 under stress controlled, fully reversed ($R=-1$), tension-compression loading clearly indicate that the 10^7 cycles fatigue limit is significantly greater in the superelastic austenite compared with thermal martensite. Furthermore, the fatigue limits in both conditions correspond to a stress amplitude in the linear elastic range at $\sim 80\%$ of the plateau stress (see horizontal dashes on the curves in Fig. 2). MM1 rationalised this observation of a fatigue limit below the stress plateau in terms of crack nucleation behaviour. Specifically, they suggest that the initial deformation mechanisms in the linear elastic region for the thermally formed martensite ($A_f=110^\circ\text{C}$)



6 Diagram of the tension-tension stress based fatigue curves to $\sim 10^6$ cycles for nominal compositions of $\text{Ni}_{50.8}\text{Ti}_{49.2}$ with an A_f of 27°C . The 10^6 cycles fatigue limit at 30°C is lower than that obtained at 60 or 160°C . Runout conditions are noted with open symbols. The corresponding critical stress levels to induce martensite are marked with arrows for data at 30 and 60°C . For these conditions, the 10^6 fatigue limit is $\sim 67\%$ of the martensite plateau stress. After Miyazaki *et al.*²²

are elastic processes. Cyclic stresses above the plateau involve martensite detwinning and martensite deformation as well as plasticity. In contrast, in the superelastic austenite ($A_f=10^\circ\text{C}$) stress induced martensitic transformations involve some degree of plasticity upon the formation of the martensite, especially in these fully annealed conditions. Close inspection of the monotonic stress-strain curves in Fig. 2 clearly illustrates the differences in superelastic austenite and thermal martensite material fatigue conditions. For example, a stress amplitude of 400 MPa (the 10^7 cycles fatigue limit for austenite) results in strains of $\sim 2\%$ in the superelastic austenite compared with over 9% strain in thermal martensite. Consequently, it is not surprising that the fatigue lifetime for superelastic Nitinol is significantly greater than that of thermal martensite. This will be discussed in greater detail in the section on 'Non-zero mean strain conditions' with respect to the effects of mean strain and strain amplitude.

There exists a narrow temperature range, between M_s and A_s , where martensite and austenite coexist (at zero applied stress) in the material. When Nitinol is in that state, austenite forms stress induced martensite on initial loading which is then stable throughout the remainder of the test. These martensite variants will be preferentially formed according to the minimum Taylor factor^{20,21} that provides maximum strain accommodation and minimum defect formation in the direction of the cyclic applied stress; consequently, for any given stress state, there will be many fewer martensite variants formed than with thermal martensite. As such, under stress controlled conditions, Nitinol with thermal martensite appears to have poorer fatigue resistance compared with material tested between M_s and A_s . Furthermore, superelastic Nitinol would be expected to have the highest fatigue limit because it requires much higher stresses to induce the martensite. In addition, the stress induced martensite will be a preferred variant that can best accommodate cyclic strains while inhibiting crack nucleation.



7 Cyclic stress–strain curves obtained at room temperature under constant strain ($\pm 2.5\%$) tension–compression conditions for Nitinol with A_f of about $+70^\circ\text{C}$, i.e. in the martensitic condition. The numbers above the curves refer to the cycle numbers. After MM1¹³

Non-zero mean stress conditions ($R=0$)

Miyazaki *et al.*²² studied stress controlled tension–tension fatigue of dogbone shaped specimens machined from $\text{Ni}_{50.8}\text{Ti}_{49.2}$ sheet that was cold rolled and aged at 400°C for 1 h to give an A_f of 27°C . Figure 6 shows their stress controlled fatigue data as a function of test temperature where the cycling conditions were from zero to maximum stress; i.e. the mean stress equals the stress amplitude. Testing at 30 and 60°C produced superelastic conditions whereas tests at 160°C were close to the M_d temperature⁷ so that material in this condition was stable (non-transforming) austenite. For this latter case, plasticity is the primary accommodation mechanism at the higher stress amplitudes, as discussed with respect to Fig. 1d. Although there are only limited data in this study, it is apparent that the 10^6 cycles fatigue limit increased with increasing test temperature. Miyazaki *et al.* suggested that this was due to the correspondence of the fatigue limit with the critical stress to form stress induced martensite, following the well known Clausius–Clapeyron relationship[†].^{3,24} It is also well known that the effective modulus increases with increasing test temperature for a given A_f temperature,²⁴ which also inversely affects the strain accumulation during stress controlled fatigue. For the fatigue tests in Fig. 6, the 10^6 cycles fatigue limit was $\sim 67\%$ of the critical stress for stress induced martensite.²² These non-zero mean stress–life fatigue results thus follow similar trends to those shown by MM1 and MM2 for fully reversed ($R=-1$) testing.

Strain–life fatigue

The majority of Nitinol fatigue investigations in the literature have been conducted under constant deformation (strain) conditions. The preference of strain based rather than stress based total life tests for Nitinol is obvious with a simple inspection of the monotonic stress–strain curve for fully martensitic or superelastic compositions (*see* Fig. 1b and c). These curves indicate that stress is linearly proportional to strain only in the (nominally) linear–elastic region. Consequently, the global strain is continuous whereas stress tends to

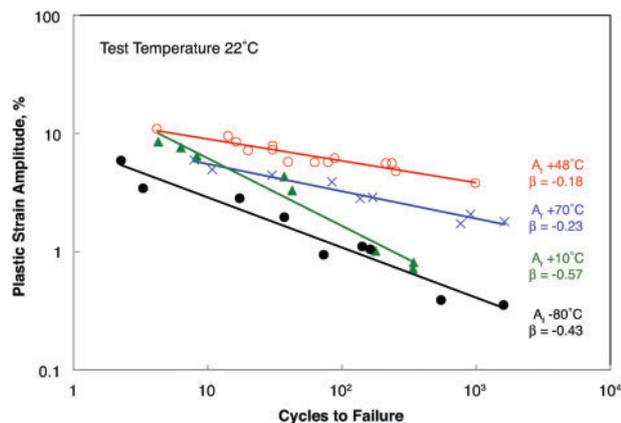
[†]As such, the stress controlled fatigue limit may also follow a constitutive equation related to the Clausius–Clapeyron relationship between $\Delta\sigma$ and ΔT , where $\Delta T=(A_f-T)$ is defined as the difference between A_f and the test temperature T .²³

‘plateau’ due to the detwinning mechanism in martensite or due to stress induced martensitic transformation in superelastic material. Therefore, in theory, the stress is constant over a wide range of (global) strains. This makes stress based test conditions less certain than strain based. Furthermore, from the perspective of experimental set-up, deformation based testing requires less sophisticated equipment. For example, rotary bending fatigue data may be obtained by rotating bars, wires or tubes around a constant radius of curvature. These strain based tests are divided into two categories: zero mean strain (where intrinsic mechanical property asymmetries play a role), and non-zero mean strain conditions (where microstructural and phase effects are more easily isolated).

Zero mean strain conditions – uniaxial fatigue

Melton and Mercier^{13,14,25} presented some of the earliest strain based low cycle fatigue data using uniaxial tension–compression specimens similar to those used for their stress controlled fatigue tests described above. Cyclic stress–strain loops are shown in Fig. 7 for the annealed $A_f \sim 70^\circ\text{C}$ thermal martensite that was cycled between strain limits of $\pm 2.5\%$ at room temperature. The material cyclically hardened, whereby the peak stress at maximum strain increased with increasing number of cycles. MM1 showed that cyclic hardening in these specimens increased rapidly over the first ~ 10 cycles; thereafter, the stress increased less rapidly, leading to cyclic stabilisation. Moreover, the curves displayed a definite anisotropy in tension versus compression properties; i.e. the absolute value of the stress at the first cycle of $+2.5\%$ strain was clearly lower than that at -2.5% strain. Differences in tension–compression mechanical behaviour in engineering materials are attributed to the anisotropy of the elastic constants.²⁶ More specifically for Nitinol, the phenomenological theory of martensitic transformations has been used effectively to model tension–compression asymmetry and the orientation dependence of the deformation behaviour.^{27–30} These analyses show that the martensite variants created in tension are different from those created in compression, which give rise to differences in plateau stresses, shape memory recovery strains and critical resolved shear stresses. Furthermore, recent detailed synchrotron X-ray diffraction analysis shows a dramatic crystallographic path difference for stress induced martensite transformations in superelastic austenite.³¹ Additional cycles at these strain limits also tended to decrease the hysteresis loop; for example, compare the 1017 cycles loop to the original loop in Fig. 7. Similar observations were made for cycling at $\pm 4.5\%$ strain. MM1 proposed that the cyclic hardening and narrowing of the hysteresis loop was likely due to the accumulation of dislocations, similar to their explanation for stress controlled fatigue conditions, discussed previously. Additional discussions of microstructural effects will be presented in the section on ‘Non-zero mean strain conditions’.

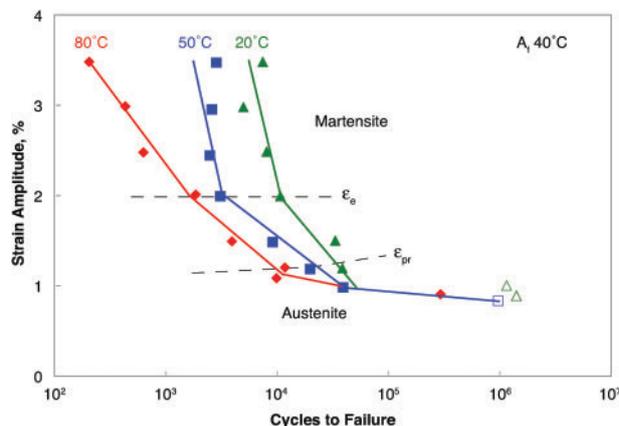
The plastic strain amplitude per cycle from the above strain controlled tests is plotted in Fig. 8 as a function of the number of cycles to failure in the low cycle regime. All graphs show a linear relation for A_f values of -80°C (stable austenite), $+10^\circ\text{C}$ (superelastic austenite), $+48^\circ\text{C}$ (coexisting austenite and thermal martensite) and $+70^\circ\text{C}$ (thermal martensite). MM1 defined the ‘plastic’ strain per



8 'Plastic' strain amplitude per cycle for four Nitinol compositions tested under constant strain tension–compression conditions at room temperature. Plastic strain amplitude is defined here as the hysteresis loop half width at zero stress. The data are fit with a least squares relationship and the resultant slope is recorded. In general, these data show that fatigue life tends to increase with increasing A_f temperature for annealed Nitinol. After Melton and Mercier^{13,14}

cycle as the total width of the stress–strain hysteresis loop at zero stress; i.e. this is the unresolved strain per cycle whether due to actual plasticity or shape memory (thermal martensite) effects. Specimens with A_f of $\sim 48^\circ\text{C}$ were tested under special conditions whereby the deformation temperature is between A_s and A_f . For this combination of transformation temperature and test temperature, stress induced martensite is formed on the first loading and then remains stable throughout the remainder of the test. As such, it appears that this stabilised martensite leads to superior low cycle fatigue behaviour since minimal damage accumulation occurs. In contrast, stable austenite ($A_f \approx -80^\circ\text{C}$), which is non-transforming at room temperature and exhibits elasto-plastic behaviour, had the lowest plastic strain amplitude of the specimens tested due to higher plastic damage accumulation. These low cycle fatigue data can be well represented by the Coffin–Manson law^{5,9,10}. Slopes of β of about -0.2 were observed for the A_f of $\sim 48^\circ\text{C}$ (stress induced martensite plus thermal martensite) and A_f of about $+70^\circ\text{C}$ (thermal martensite) data, implying similar deformation mechanisms in the two alloys. In contrast, the corresponding slope was -0.43 from the stable austenite alloy ($A_f \approx -80^\circ\text{C}$), which is more similar to the approximately -0.5 slope from traditional engineering materials.^{9,10,32} For the superelastic A_f of about $+10^\circ\text{C}$ alloy with $\beta = -0.57$, MM2 speculated that there is an interaction between the stress induced martensite and the dislocation debris that forms with each fatigue cycle. These data demonstrate that under strain controlled conditions larger cyclic strains to fracture are required at 10^3 cycles as compared with other engineering materials. As noted by MM1,¹³ it is remarkable that a fatigue life of 1000 cycles is obtained for a cyclical strain amplitude of 10%, where conventional engineering materials sustain only a strain amplitude of $<1\%$ before comparable cyclic fracture. Considering these data for low cycle fatigue,

⁵ The Coffin–Manson equation is an empirical relationship for low cycle fatigue relating the number of cycles to failure N_f to the (plastic) strain amplitude per cycle $\Delta\epsilon_p$, namely

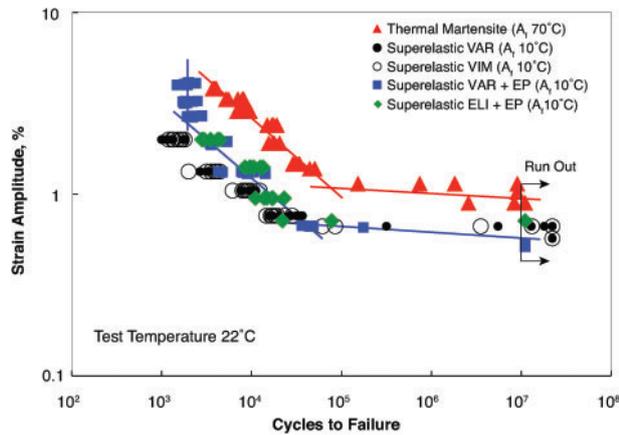


9 Effect of cyclic strain amplitude on the rotary bending fatigue life for $\text{Ni}_{50.9}\text{Ti}_{49.1}$ wires at three test temperatures for an A_f of 40°C as determined by DSC. The open symbols indicate runouts for these test conditions. The low cycle fatigue behaviour is greatly affected by the test temperature with an order of magnitude improvement in fatigue at 20°C compared with that at 80°C . However, with this limited dataset, there does not appear to be a difference at the high cycle regime. After Kim and Miyazaki³³

larger total strains for given lifetime are achieved when the deformation temperature is between A_s and A_f compared with deformation of thermally formed martensite for these fully annealed materials. This observation is consistent with the argument presented above that fewer stress induced martensite variants are formed and that these variants are then more mobile on stress reversal.

Zero mean strain conditions

Rotary bending tests are a convenient way to collect strain based fatigue data for wire specimens under zero mean conditions. Strain amplitude is approximated by the ratio of wire diameter to curvature diameter, so a range of strain amplitudes can be achieved for a given wire diameter with slight adjustments in the curvature. Test temperature can also be easily adjusted with controlled environment baths. Miyazaki and co-workers studied the effects of strain amplitude and temperature on fatigue life on $\text{Ni}_{50.9}\text{Ti}_{49.1}$ wires out to 10^6 cycles.^{33,34} The wires were cold drawn 30% and then aged at 400°C for 1 h, resulting in an A_f temperature of 40°C [measured via differential scanning calorimetry (DSC)]. Fatigue lives, measured at test temperatures between 20 and 80°C , are shown in Fig. 9 in terms of the strain amplitude versus the number of rotations to fracture. There is a general trend of increasing fatigue life with decreasing test temperature in the high cycle fatigue regions (high and intermediate strain amplitude regions). For these limited data, the fatigue strain at 10^6 cycles appears to be insensitive to temperature. The authors characterise the fatigue behaviour with respect to the phases present (superelastic austenite and thermal martensite) and the transitions between ϵ_{pr} (proportional strain limit) and ϵ_e (elastic strain limit) from the monotonic stress–strain curves at the respective test temperatures. A key teaching from this study is that in the short life regime, there is cyclic austenite–martensite or martensite–martensite variant transformation and in



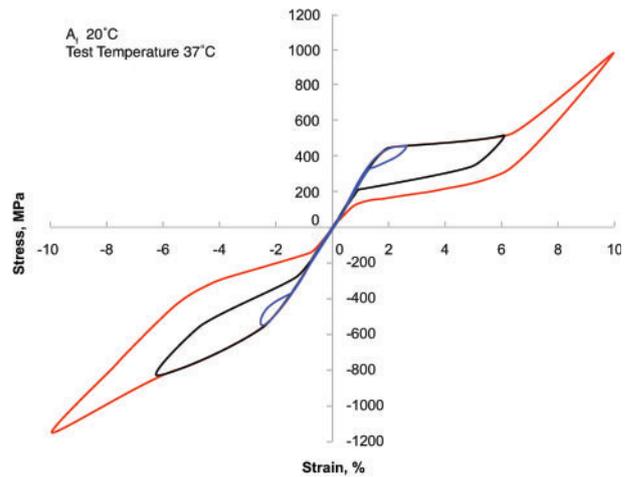
10 Comparison of rotary bending fatigue data at room temperature from $\text{Ni}_{50.8}\text{Ti}_{49.2}$ (superelastic with $A_f \approx 10^\circ\text{C}$)^{35,36,38} and $\text{Ni}_{49.5}\text{Ti}_{50.5}$ (thermal martensite with $A_f = 70^\circ\text{C}$)³⁷ thermomechanically treated wires. The thermal martensite data have greater low cycle and high cycle lives compared with the three sets of data for superelastic conditions. Note the three distinct regions for the superelastic data, whereby strain amplitudes greater than 2.5% have a constant live of ~ 1800 cycles, corresponding to cycling on the stress plateau. The four sets of superelastic data compare melting practice [vacuum arc remelt (VAR), extra low impurity and vacuum induction melting (VIM)]^{35,38} as well as surface finish (black oxide³⁵ and electropolished^{36,38}). See text for additional discussion. Specimens that reached $\geq 10^7$ were considered as runouts, as illustrated by the arrows

the long life regime, there is a constant phase fraction with only elastic deformations.

A comparison of rotary bending data from continuously thermomechanically processed wires (per modern practice) is shown in Fig. 10 for superelastic and thermal martensite conditions out to 10^7 cycles tested at room temperature.^{35–38} The four sets of superelastic data are consistent and have shorter lives at equivalent strain amplitudes compared with the thermal martensite data up to 4.3% strain amplitude. These observations can be explained by examining the corresponding fully reversed stress–strain curves (see e.g. Fig. 11). For a given strain amplitude, the stress required to deform thermal martensite is lower than those required to stress induce martensite under superelastic conditions. Furthermore, it has been shown that the crack growth rates for martensite and superelastic materials are virtually identical (see damage tolerant fatigue section); therefore, under strain controlled conditions, it appears that the crack initiation rates are faster in the superelastic condition than in thermal martensite.

Close inspection of Fig. 10 suggests that the superelastic data can be divided into three distinct regions:

- (i) low cycle fatigue between tension and compression of stress induced martensite with a constant ~ 1800 cycles to failure. This strain amplitude range represents cycling on the stress plateau; see, for example, Fig. 11 for an idealised uniaxial tension–compression curve for superelastic Nitinol. Consequently, the cyclic deformation mechanism of superelastic Nitinol versus thermal martensite is similar in this strain range. Note



11 Schematic of fully reversed stress–strain data for superelastic Nitinol under bending conditions illustrating the stress hysteresis within the linear–elastic range, in the combination of linear and stress induced martensite, and cycling on the stress plateau

that the thermal martensite data do not show this constant fatigue life in this strain range since there is no stress induced phase transformation

- (ii) intermediate cycle fatigue between $\sim 10^3$ and 10^4 cycles with an increasing life with decreasing strain amplitude. This region corresponds to the transition between linear–elastic and superelastic behaviour; from Fig. 11, the stress hysteresis in this cyclic range is significantly lower than that for the $\pm 4\%$ cyclic behaviour. This range corresponds to the Coffin–Manson region for fatigue behaviour of engineering materials, as discussed above. However, the empirical Coffin–Manson analysis^{9,10} is explicitly ascribed to cyclic plastic strain. As such, it is not appropriate to make direct comparisons with cyclic superelastic data since a reverse phase transformation is part of the deformation accommodation. Nevertheless, without implying similar mechanisms, it is interesting to characterise the slope of the log–log best fit curve of -0.41 for the Sheriff *et al.*³⁶ data and -0.34 for the Reinoehl *et al.*³⁵ data. The difference between the two datasets may be due to the slight difference in the experimental set-up, where Sheriff used a fixed diameter mandrel to hold the strain constant. On the other hand, the Reinoehl wire specimens were set to the initial curvature and then allowed to adopt slightly different curvatures throughout the test[†]. In comparison, the thermal martensite data had a slope of -0.44 . For this range of strain amplitudes, the deformation mechanism for the superelastic wires involves a combination of linear–elastic and reversible stress induced martensite phase transformation. For the thermal martensite, the mechanism involves linear elasticity and deformation of twinned martensite. Although these mechanisms are different, both sets of materials were processed with $\sim 40\%$ cold work and

[†] Berg³⁹ and Wick *et al.*⁴⁰ showed that superelastic Nitinol does not create a smooth curvature under a constant moment bend. This is due to the non uniformity of stress induced martensite formed.

then stress relieved continuously at $\sim 500^\circ\text{C}$.⁷ Consequently, the starting microstructures for the superelastic and thermal martensite contain a large density of dislocations, fine grain size ($\leq 1\ \mu\text{m}$) that help to stabilise the fatigue resistance

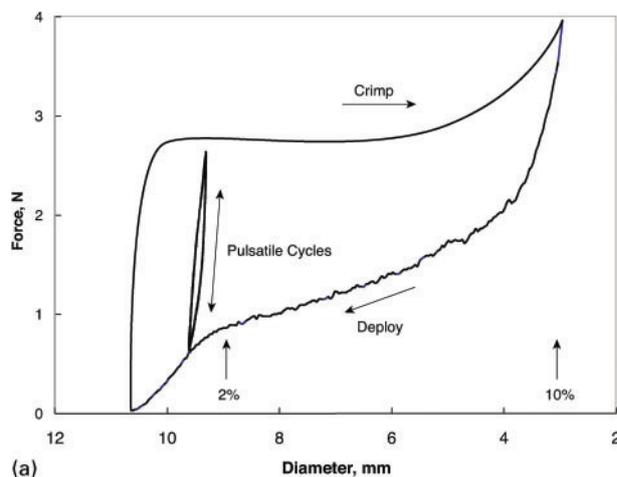
- (iii) high cycle fatigue represents linear–elastic behaviour, where it is assumed that fatigue mechanism occurs due to typical damage accumulation in this (nominally) linear–elastic region, consistent with Figs. 5, 6 and 9.⁴¹

The data in Fig. 10 can also be inspected to determine the effects of melting practice and surface finish. The two sets of Reinhoehl data compare the fatigue behaviour of 0.267 mm diameter superelastic wires produced from VAR versus VIM ingots.³⁵ The VIM ingots are melted in a graphite crucible that typically adds up to ~ 500 wppm carbon to the composition; these materials tend to contain both oxide inclusions ($\text{Ti}_4\text{Ni}_2\text{O}_x$) as well as TiC precipitates. The VAR produced material generally has lower overall interstitial content (< 50 wppm C), with no TiC, but may contain slightly larger oxide inclusions; both materials are consistent with composition and microstructure per ASTM F2063-05.⁴² Reinhoehl *et al.* found that both VIM and VAR Nitinol wires had equivalent fatigue behaviour within the strain amplitude and range of cycles, regardless of the differences in carbon content and inclusion size. As an additional comparison, the data from Morgan *et al.*³⁸ from VAR+extra low impurity had an oxygen content of 90 wppm, as compared with ~ 300 wppm for VAR and VIM+VAR compositions. Note that even with a lower oxygen content that led to a substantially lower inclusion volume fraction, the rotary bending fatigue data are indistinguishable from the more standard melt practices. This suggests that within the limits of composition of the ASTM standard and for up to 10^7 fatigue cycles, the size of these inclusions is below the critical crack size for failure (about 15–50 μm),⁴³ thereby producing the same total fatigue life under these test conditions.

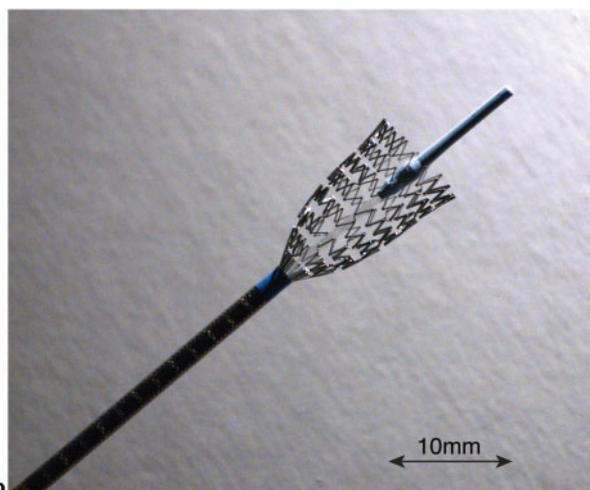
The effects of surface finish on the fatigue of Nitinol can be evaluated by comparing the results of Reinhoehl *et al.*³⁵ for black oxide wires with those of Sheriff *et al.*³⁶ for electropolished samples with similar processing conditions and resultant A_f temperatures. Electropolishing is an electrochemical process that removes surface oxides (and other surface contaminants, such as lubrication) and creates a smooth and bright surface. In the low cycle fatigue regime ($< 10^5$ cycles), electropolishing resulted in a slight, but statistically significant, increase in fatigue life; however, there is no statistically significant difference in the high cycle fatigue regime.³⁷

Non-zero mean strain conditions

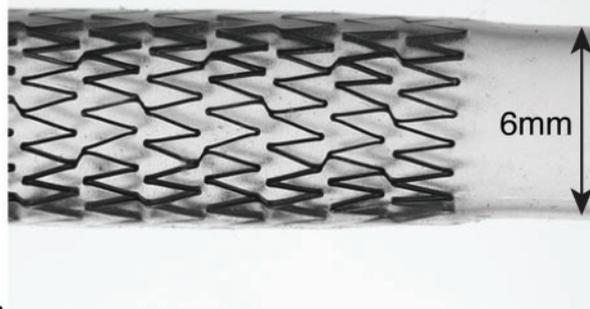
Much of the early work on Nitinol fatigue was conducted at fully reversed stresses or strains, i.e. $R = -1$. However, many modern applications of Nitinol based devices present an entirely different deformation profile and require a different method to characterise the fatigue conditions. For example, in medical devices such as endovascular self-expanding stents, the Nitinol device may undergo a large single strain excursion (up to 10%) during constraint into a delivery system and then the opposite strain excursion during deployment into the anatomical location.^{44–46} Upon implantation, the Nitinol devices are slightly oversized with respect to the



(a)



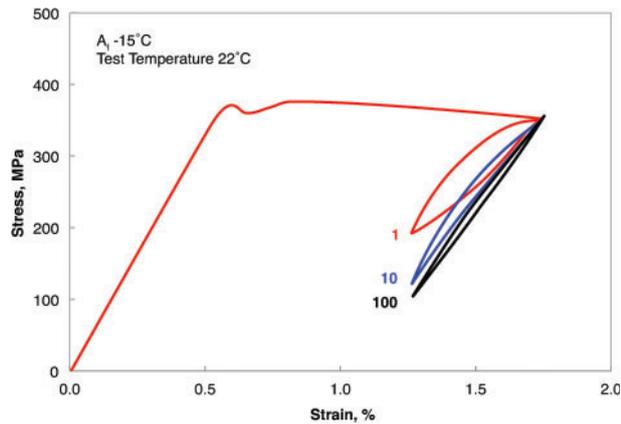
b



c

- 12 **a** stress–strain data for superelastic Nitinol schematically illustrating the initial crimp stress, deployment and subsequent *in vivo* cycles. **b** Cordis Nitinol SMART stent being deployed from a delivery catheter; the strain on the stent inside the catheter can be $\sim 10\%$. **c** illustration of a 9 mm diameter Cordis Nitinol SMART stent in a 6 mm diameter mock vessel, which simulates a constant *in vivo* mean strain. Reprinted from Pelton *et al.*⁴⁶ with permission from Elsevier

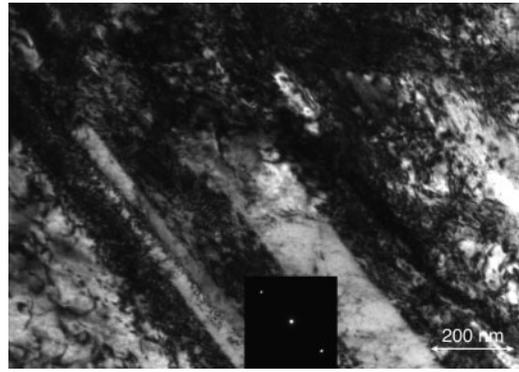
anatomical locations. This ‘interference fit’ with the anatomy produces conditions of a constant mean strain on the Nitinol device with lower strain amplitudes; this condition is schematically shown in Fig. 12a. The physiological movements from the cardiac systolic–diastolic cycle and musculoskeletal motion provide the imposed strain amplitude, as illustrated in the pulsatile fatigue cycle in Fig. 12a. Figure 12b and c shows a Cordis



13 Stress–strain fatigue curves from annealed $\text{Ni}_{50.8}\text{Ti}_{49.2}$ bar under conditions of 1.5% mean strain and strain amplitude of 0.25%. Cycles 1, 10 and 100 are shown whereby the hysteresis width decreases, the modulus increases, with an effective increase in the stress range with increasing cycles. Reprinted from Pelton⁴⁷ with permission from Springer

Nitinol SMART stent with partial deployment from a delivery system and as deployed in a mock arterial vessel respectively.⁴⁶ Owing to this complex application of strains, biomedical applications of Nitinol are often quite complicated with respect to expectations of fatigue durability; the full range of ‘duty cycles’ must therefore be addressed in a testing and analysis programme in order to illuminate fatigue safety. The key to understanding Nitinol based medical device fatigue and durability is to establish the Nitinol fatigue properties that are mainly material and surface condition dependent and the device durability properties that are design and material dependent.

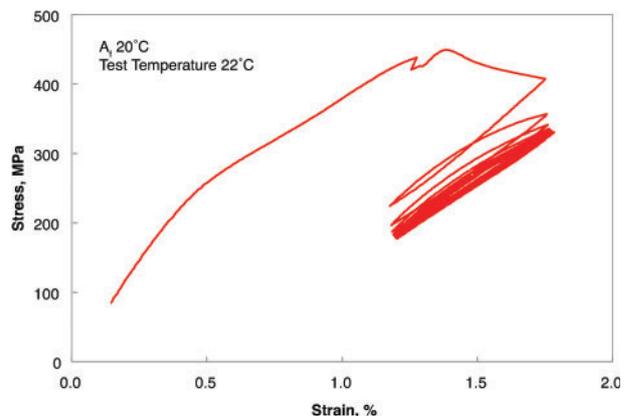
As discussed above, many of the early fatigue studies on Nitinol were conducted on fully annealed material with grain sizes of $\sim 100\ \mu\text{m}$ and with low defect density. These studies, including the suite of seminal Melton and Mercier papers, provide insight into many of the fundamental mechanisms. However, they are incomplete with respect to providing a definitive microstructure mechanics explanation that could be used to predict fatigue behaviour of modern medical devices that exclusively use thermomechanically processed Nitinol material. In an effort to understand the effects of processing, mean strain and strain amplitude, and elucidate a concomitant microstructural understanding, Pelton conducted a transmission electron microscopy (TEM) investigation on fatigued superelastic $\text{Ni}_{50.8}\text{Ti}_{49.2}$ specimens.⁴⁷ Tensile specimens were machined from superelastic wire with a 3 mm diameter gauge. The specimens were either annealed at 900°C for 30 min ($A_f = -15^\circ\text{C}$) or stress relieved at 500°C for 10 min ($A_f = 20^\circ\text{C}$). Figure 13 shows the strain controlled fatigue curves for the annealed specimen at 1.5% mean strain and 0.25% strain amplitude for 1, 10 and 100 cycles tested at room temperature. With increasing cycles, the hysteresis width decreased, the modulus increased to a value comparable with the initial austenite modulus, with a concomitant increase in stress range; these changes are comparable with those discussed by MM1 for $R = -1$ test conditions. The original annealed microstructure consisted of equiaxed grains of approximately $50\text{--}100\ \mu\text{m}$



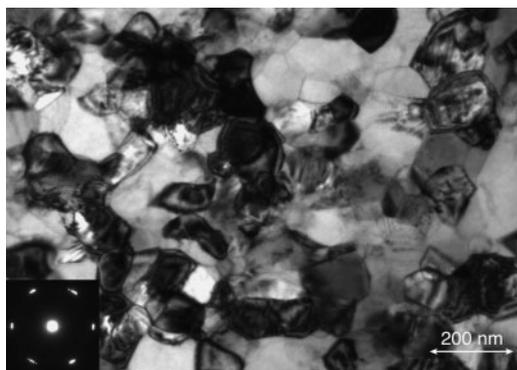
14 Transmission electron microscopy microstructure in annealed Nitinol after 10 cycles consists of banded austenite with a high density of dislocations ($\sim 10^{14}\ \text{m}^{-2}$) with $a\langle 100\rangle\{011\}$ slip systems interspersed with retained deformed martensite. Reprinted from Pelton⁴⁷ with permission from Springer

with low dislocation density ($\sim 10^8\ \text{m}^{-2}$); these dislocations are generally in proximity to grain boundaries or $\text{Ti}_4\text{Ni}_2\text{O}_x$ particles. In stark contrast, the TEM microstructure after only 10 fatigue cycles consists of banded austenite with a high density ($\sim 10^{14}\ \text{m}^{-2}$) of dislocations with $a\langle 100\rangle\{011\}$ slip systems interspersed with retained deformed martensite (Fig. 14). An increase in dislocation content with increasing cycles was also observed for thermally cycled annealed Nitinol.⁴⁷ The observations of dislocation bands suggest that the moving martensite interfaces act as crystallographic ‘snowplows’ that force the new and existing $\langle 100\rangle\{011\}$ shear loops to slip (and cross-slip) and then become locked in the sessile bands at the interface plane.⁴⁷ Although no TEM investigations were conducted during the MM studies discussed above, it is likely that similar microstructural effects were responsible for the observed changes in hysteresis width and peak stress after even more severe $R = -1$ cyclic testing.

Fatigue curves that correspond to the stress relieved superelastic condition are shown in Fig. 15.⁴⁷ Note that this optimised processing (typical of Nitinol medical device manufacturing) results in a more stabilised fatigue behaviour with minimal softening and a relatively stable cyclic modulus that is comparable with the



15 Stress–strain fatigue curves from stress relieved superelastic $\text{Ni}_{50.8}\text{Ti}_{49.2}$ under conditions of 1.5% mean strain and strain amplitude of 0.25%. Note that the cyclic curves follow the R phase slope and are more stable than those shown in Fig. 13. Reprinted from Pelton⁴⁷ with permission from Springer



16 Transmission electron microscopy microstructure in stress-relieved Nitinol after 10 cycles consists of equiaxed 75 nm austenite grains with a $\langle 111 \rangle$ orientation, which produces more stable fatigue conditions, as shown in Fig. 15. Compare with the banded and dislocated microstructure after cycling the annealed structure (Fig. 14). Reprinted from Pelton⁴⁷ with permission from Springer

initial stress induced R phase modulus. In contrast to the annealed microstructures, the stress-relieved and 10 cycles fatigued microstructure has an equiaxed grain size of ~ 75 nm with a $\{111\}_A$ texture (see inset diffraction pattern) as shown in Fig. 16. These austenitic grains contain dislocations and subgrains (i.e. grains with low angle grain boundary separation) with no apparent retained martensite, similar to the non-fatigued microstructures.

Based on the TEM results discussed above,⁴⁷ it is clear that processing plays a key role in the fatigue behaviour of Nitinol. Several authors emphasised the importance of combining cold work and aging treatments to optimise the microstructure of Nitinol for cyclic stability.^{3,16,17,48} The amount of unresolved strain as a result of cycling is a quantitative measure related to shape retention; it pertains to the cyclic damage accumulation. In general, a cold work plus age treatment produces a more stabilised microstructure to minimise unresolved strain than the anneal plus age treatment for cyclic applications, as demonstrated in Figs. 13–16.

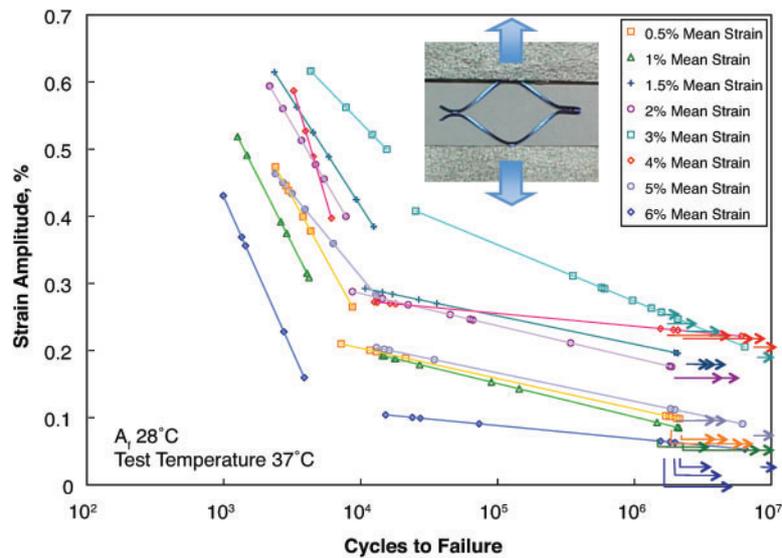
Since these treatment parameters vary from product to product, testing under simulated *in vivo* conditions using process-optimised Nitinol is then required to gain insight into the device-specific fatigue behaviour. Several studies have emerged recently to address the fatigue behaviour of thermomechanically processed Nitinol under conditions of a constant mean strain and strain amplitude.^{46,49–53} The results of these investigations truly demonstrate the unique fatigue properties of Nitinol, as compared with traditional engineering materials.

Tabanli *et al.*^{49,50} were the first to observe non-traditional constant life fatigue behaviour on super-elastic Nitinol microtubing with an A_f of about -5°C . Their room temperature experiments were performed at a frequency of 20 Hz and involved cycling to failure at a strain amplitude of 0.22% with tensile mean strains from 0.25 to 9.31%. Five conditions of mean strain and strain amplitude were chosen to span the monotonic stress–strain space to include deformation of linear–elastic austenite, mixed phases of austenite and martensite (i.e. superelastic) and linear–elastic martensite (i.e. beyond the upper plateau stress but below plastic deformation strains), as summarised in Table 1. With only five test specimens, the data are not statistically significant; nevertheless, these results are quite revealing. The longest lives were obtained under conditions of cyclic deformation in linear–elastic austenite (mean strain = 0.25 and 0.36%) and superelastic material that had been loaded high enough to stabilise the stress-induced martensite (mean strain = 9.3%). Mixed phase cycling in the superelastic range (i.e. on the upper plateau) produced shorter lives in these specimens; again, this is consistent with the conclusions drawn by Miyazaki and Wick, discussed above. Tabanli *et al.*^{49,50} concluded that the typical Goodman and Soderberg constructions^{41,54} for a constant life diagram are not valid for Nitinol since there was no decrease in fatigue life with increasing mean strain. Specifically, they hypothesised that a long fatigue life requires no cyclic phase transformations and provided a mean alternating strain fatigue life diagram based on that assumption. It is also interesting that the fatigue lives for these specimens were all well less than 10^6 cycles. There are two possible explanations when comparing these results with other non-zero mean strain fatigue investigations.^{46,51–53} First, the test temperature for the Tabanli experiments was $\sim 30^\circ\text{C}$ greater than the A_f (ΔT of $\sim 30^\circ\text{C}$). As discussed previously for zero mean strain conditions, the fatigue life in Nitinol tends to increase as the difference between the test temperature and A_f decreases (i.e. with a smaller ΔT). Second, Macherauch⁵⁵ summarised the general effects of deformation mode on fatigue life of engineering materials and showed that tension–tension conditions lead to the shortest lives compared with bending fatigue or rotary beam fatigue test conditions. The Nitinol tubes in the Tabanli investigation were tested under uniaxial fatigue conditions, whereby the entire cross-section of the tubing was exposed to the cyclic strains. Conversely, in rotary beam fatigue, each section of the test specimen is alternately in tension and compression with a strain gradient across the section diameter. Therefore, the differences in the data can be ascribed to the test technique chosen, whereby the active

Table 1 Mean strain effects on fatigue life in Nitinol; after Tabanli *et al.*⁴⁹

Mean strain/%	Strain amplitude/%	Cycles to fracture	Condition
0.25	0.21	185 000	Linear–elastic austenite
0.36	0.20	97 000	Linear–elastic austenite
2.12	0.24	10 900	Superelastic (loading plateau)
2.0*	0.22	24 340	Superelastic (unloading plateau)
9.3	0.22	82 650	Deformed martensite

*Specimen was prestrained to 9.5% strain and then unloaded to 2.0% before cycling to simulate conditions involving the implantation of a Nitinol medical device.

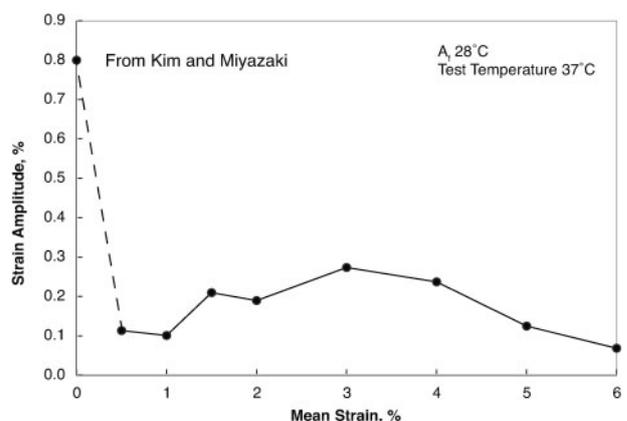


17 Effects of mean strain and strain amplitude on the cycles to fracture for Nitinol stent subcomponents. The test specimen and affixed grip fixture are shown in the inset along with arrows to illustrate the direction of cyclic deformation. Testing was conducted at 37°C at 25 Hz, whereby specimens were first loaded to 6% prestrain and then unloaded to the mean strain and cycled on the unloading plateau. Regression analyses were used to determine the best fit for the low cycle and high cycle regions. After Tolomeo *et al.*⁵¹

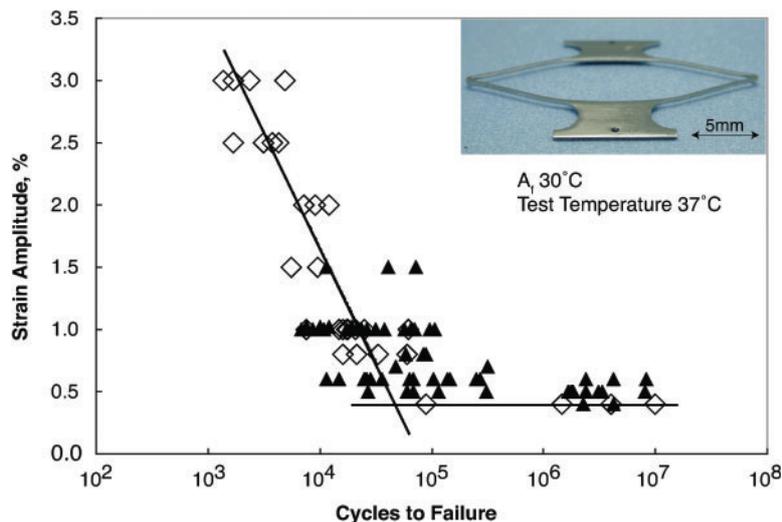
volume and/or surface subjected to maximum principal strains was effectively higher in the uniaxial tensile specimens.

Tolomeo *et al.*⁵¹ investigated the fatigue behaviour of fully processed Nitinol endovascular stent components (including laser machining, expansion and processing) under several conditions of mean strain and strain amplitude. For these experiments, subcomponents were extracted from the stents that were cycled about mean and amplitude displacements. An image of the stent subcomponent is inset in the graph in Fig. 17. Non-linear finite element analysis (FEA) was used to convert the global displacements of the stent to local strains (*see, e.g. Rebelo et al.*),^{56,57} From a medical device perspective, the advantage of this type of fatigue specimen is that the subcomponents are processed exactly as the medical device, unlike conventional ‘dogbone’, tube or wire specimens that are obtained from raw materials. Consequently, these specimens reflect the actual material conditions of the device, *e.g.* they contain typically preexisting surface defects so that both crack initiation and propagation are evaluated during the test. To avoid possible testing artefacts (such as out of plane buckling), from physiologically relevant compression (*i.e.* ‘push’ conditions), testing for these specimens was conducted under tension–tension (‘pull’) conditions. As such, the maximum alternating strains were on the inside of the stent strut rather than the outside as experienced by stents subjected to *in vivo* biomechanical deformations due to radial pulsatile motion. Figure 17 shows the ϵ_a – N diagram for the Tolomeo tests for specimens that were initially strained to 6% (prestrain) and then unloaded to mean strains ranging between 0.5 and 6% before being cycled with strain amplitudes between 0.05 and 0.6%; *i.e.* testing was conducted with mean strain on the unloading plateau. Note that for a given strain amplitude, those specimens that were subjected to 3% mean strain resulted in the longest lives in both the low cycle and high cycle regions. In contrast, application of 1 and 6% mean strains resulted in the shortest lives for a given strain amplitude.

The mean and strain amplitudes to give a constant 10^6 cycles life for each condition were determined from these runout data and are plotted on a constant life diagram as shown in Fig. 18. The zero mean condition was obtained from rotary bending fatigue data from Kim and Miyazaki,³³ the validity of this selection of zero mean data is questionable due to the arguments above with respect to deformation mode. Nevertheless, the data in Fig. 18 clearly show that for a constant 10^6 cycles life, the strain amplitude is $\sim 0.1\%$ at mean strains of 1 and 6% compared with 0.3% strain amplitude at 3% mean strain. Morgan *et al.*⁵² also observed that the fatigue life of Nitinol wire ($A_f=12^\circ\text{C}$) increased with increasing mean strain from 2 to 6% for strain amplitudes that ranged from 0.5 to 3.0% in 38°C Ringer’s solution (simulated physiological solution). Their specimens were strained to the mean strain and cycled with no initial prestrain (cycling from the upper plateau). They postulated that the advancing fatigue



18 Effects of mean strain and strain amplitude on the cycles to fracture for Nitinol stent subcomponents to 10^6 cycles. The zero mean datum is after Kim and Miyazaki³³ rotary bending of Nitinol wire at 10^6 cycles. After Tolomeo *et al.*⁵¹

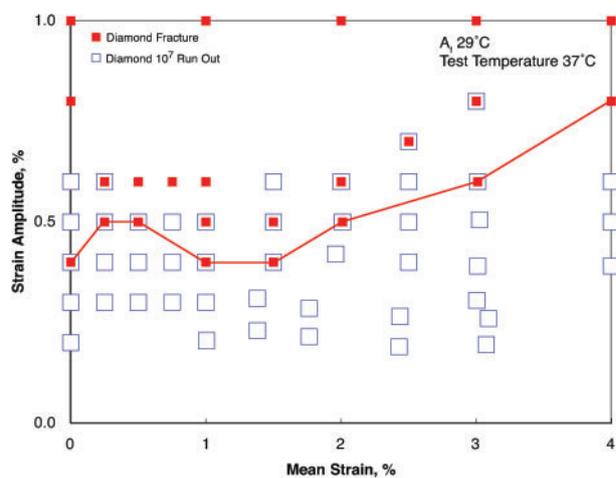


19 Strain amplitude versus fatigue data (strain-life) in Nitinol for conditions that led to fracture for zero mean strain (open symbols) and non-zero mean (closed symbols) are presented. The diagram can be separated into low cycle (10^3 – 10^5 cycles) and high cycle ($\geq 10^5$ cycles) regions. The fatigue strain limit at the high cycle data is 0.4%, corresponding to the lowest strain amplitude where fractures were observed. A Nitinol diamond shaped specimen manufactured from a laser machined microtube is inset. Specimens were thermally processed and electropolished similar to endovascular stents. Reprinted from Pelton *et al.*⁴⁶ with permission from Elsevier

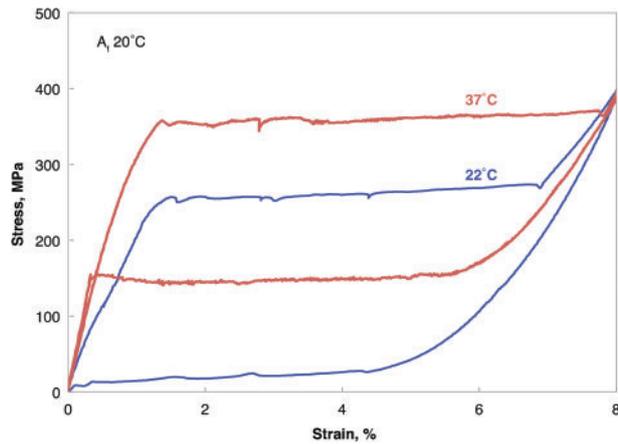
crack tip increases the stress intensity to a level where martensite is induced ahead of the crack tip and stabilised by residual stresses and plastic deformation. The results confirm both the importance of transformation strain and phase state within the material during fatigue loading. These investigations also substantiate the primary conclusion from Tabanlı *et al.*^{49,50} that the typical stress based Soderberg or Goodman type analyses are not appropriate for fatigue of superelastic Nitinol because of the stress independence from strain along the loading and unloading plateaus.

Pelton *et al.* investigated the effects of mean strain and strain amplitude on Nitinol diamond shaped test coupons cycled in bending to 10^7 cycles and compared these results with radial pulsatile fatigue testing of commercially available Nitinol self-expanding stents.^{46,53} The inset in Fig. 19 shows a diamond shaped subcomponent that was designed to simulate a representative unit cell of a commercial stent (8 mm length, 0.33 mm width and 0.35 mm thick), and hence assess the pulsatile fatigue properties of a Nitinol biomedical device. Specimens were laser machined from Ni_{50.8}Ti_{49.2} tubing (4.67 mm outside diameter with 0.38 mm wall thickness) and were processed to achieve a target A_f of 30°C, similar to commercially available Nitinol stents ($\Delta T \approx 7^\circ\text{C}$). As such, these diamond specimens represented, as closely as possible, the material properties, relative geometry and processing conditions of the actual stents. In these studies, a total of 432 specimens were cycled, either to fracture or to 10^7 cycles life, at various combinations of mean and strain amplitudes at 37°C; a minimum of four specimens were tested at each condition. Results from these fatigue studies are shown in Fig. 19. Fatigue results from test conditions ranging from 0 to 4% mean strains and cyclic strain amplitudes from 0.2 to 1.5% are shown, with the non-zero mean strain diamond fatigue data (closed symbols) superimposed on the zero mean strain data (open symbols). The diamond specimens were first subjected to ‘compression’ (push) prestrain of 5%; consequently, the fatigue cycling was on the unloading

rather than loading portion of the stress-strain curve. Note the large variation in the cycle life especially at the low values of strain amplitude, which is likely due to natural variations in fatigue behaviour as well as the effects of mean strain, as discussed below. The diamond fatigue life data from Fig. 19 are replotted in a more illuminating way as a constant life diagram as mean strain versus strain amplitude in Fig. 20. The specimens that survived 10^7 cycles are shown as open squares, whereas those specimens that fractured are shown as solid squares.



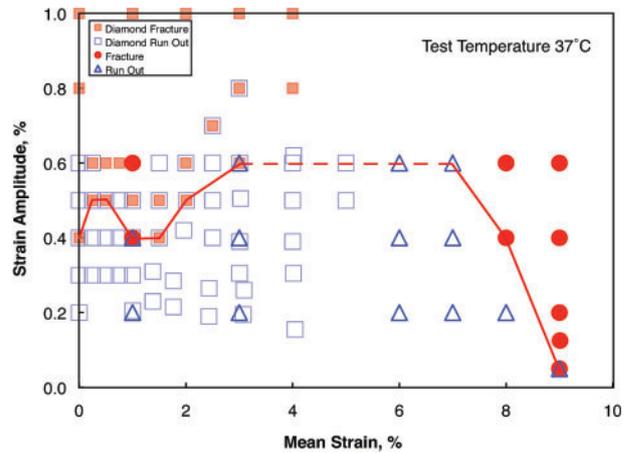
20 Constant life diagram from the diamond stent subcomponent fatigue testing where the various conditions of mean strain and strain amplitude are plotted. Conditions that survived the 10^7 cycles testing are shown as open squares, whereas cyclic conditions that led to fracture $< 10^7$ cycles are represented with closed squares. The lowest strain amplitude that resulted in fracture $< 10^7$ cycles is illustrated with a line. Note that mean strains greater than $\sim 1.5\%$ lead to greater fatigue life, in contrast to conventional linear-elastic materials. Reprinted from Pelton *et al.*⁴⁶ with permission from Elsevier



21 Monotonic stress–strain curves from Nitinol microdogbone specimens that were extracted from stent-like devices that were processed similar to Nitinol medical devices. Note that the plateau length at 37°C is ~8% in length. Reprinted from Pelton⁴⁷ with permission from Springer

The lowest strain amplitude that resulted in fracture is illustrated with a line, whereby combinations of mean strain and strain amplitude below this line should result in lifetimes of at least 10^7 cycles. These data clearly show that an increase from 1.5 to 4% mean strain (i.e. increase in volume fraction of stress induced martensite) results in a greater tolerance of strain amplitude from 0.4 to 0.8% respectively.

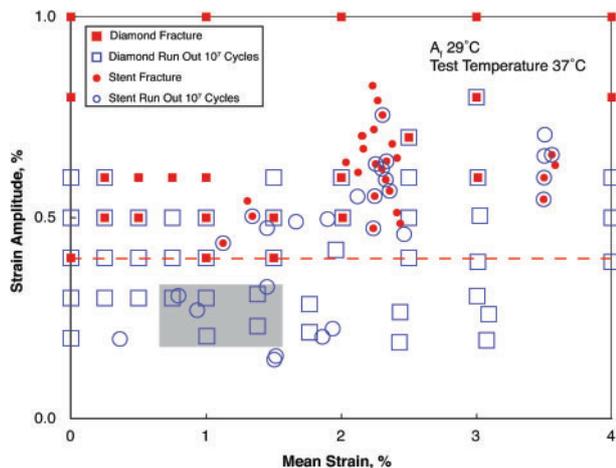
More recently, microdogbone specimens (6 mm gauge length, 0.3 mm gauge width and 0.15 mm gauge thickness) were extracted from stent-like devices that were laser machined from thermomechanically processed $\text{Ni}_{50.8}\text{Ti}_{49.2}$ tubing and expanded per standard Nitinol stent processing.⁴⁷ The monotonic stress–strain curves at room temperature and 37°C from these specimens shown in Fig. 21 illustrate the expected trend of increased plateau stresses, greater modulus and longer stress plateau when tested at 37°C. Specifically, the length of the stress plateau (linear elastic plus stress induced plateau) at 37°C is ~8% versus only 6.5% in the room temperature tested samples. The microdogbone specimens were prestrained at 37°C to 9% (beyond the stress plateau) to simulate Nitinol stent crimp strains and then fatigued from the unloading plateau. Data from three specimens per mean strain condition are superimposed on the ‘diamond’ data and are shown in Fig. 22. At 1% mean strain, the dogbone fatigue data are comparable with those from the diamond shaped specimens; i.e. no fractures at 0.2% strain amplitude, fracture at 0.6% strain amplitude and mixed results at 0.4% strain amplitude. For mean strains between 1.5 and 3%, there is an increase from 0.4 to 0.6% strain amplitude. A dotted line at 0.6% strain amplitude is drawn between 3 and 7% mean strain to indicate that there are insufficient data for a complete analysis. Above 7% mean strain (approximately the superelastic plateau length at 37°C), however, the constant life data exhibit a negative slope. These observations demonstrate that Nitinol has a worst case fatigue life with a strain amplitude of 0.4% up to a mean strain of ~1.5%. Above 1.5%, however, the strain amplitude increases for a given fatigue life, with the strong implication that the formation of stress induced martensite is responsible for



22 Constant life diagram from the diamond stent sub-component fatigue testing where the various conditions of mean strain and strain amplitude are plotted. Also included are data from microdogbone specimens laser machined from $\text{Ni}_{50.8}\text{Ti}_{49.2}$ tubing with mean strains out to 9%. Conditions that survived the 10^7 cycles testing for the microdogbone specimens are shown as open triangles, whereas cyclic conditions that led to fracture 10^7 cycles are represented with closed circles. The microdogbone data are consistent with those from diamond specimens (superimposed) from Fig. 20. Note that mean strain leads to increasing fatigue life in the 1.5–7% mean strain region, with a decrease in strain amplitude >7% mean strain (extent of the monotonic stress plateau). Reprinted from Pelton⁴⁷ with permission from Springer

this enhancement in fatigue behaviour. Note also that in the two phase region, increasing the peak maximum strain (i.e. mean strain plus strain amplitude) does not necessarily decrease the fatigue life. For example, consider the two conditions of fatigue strain: $2 \pm 1\%$ and $2.5 \pm 0.5\%$; the peak strain is 3% for both cases. However, the lower mean strain conditions lead to fracture, whereas an increase in mean strain and corresponding decrease in strain amplitude leads to 10^7 cycles survival. The overwhelming conclusion from these recent studies is that fatigue life for superelastic Nitinol is driven by strain amplitude rather than mean strain or peak strain up to the end of the stress plateau. Furthermore, Nitinol fatigue should not be modelled with traditional Goodman or Soderberg type analyses typically used with conventional engineering materials.

The utility of such data is in the prediction of component lifetimes, e.g. in the above mentioned case for the integrity of endovascular stents. Accordingly, as validation of the pulsatile fatigue behaviour, radial distension to fracture proof tests were also conducted on 10 mm diameter Cordis SMART stents.⁴⁶ Non-linear FEA was used to determine the resultant mean strains and strain amplitudes under various conditions of oversizing and pulse pressure. These experimental stent pulsatile fatigue data are overlaid on the diamond fatigue life data in Fig. 23. The stents fractured at $<10^7$ cycles at high mean and high strain amplitude conditions, as predicted by the conservative 0.4% fatigue strain line. The stents survived to 10^7 cycles at conditions of ~1% mean strain and strain amplitudes

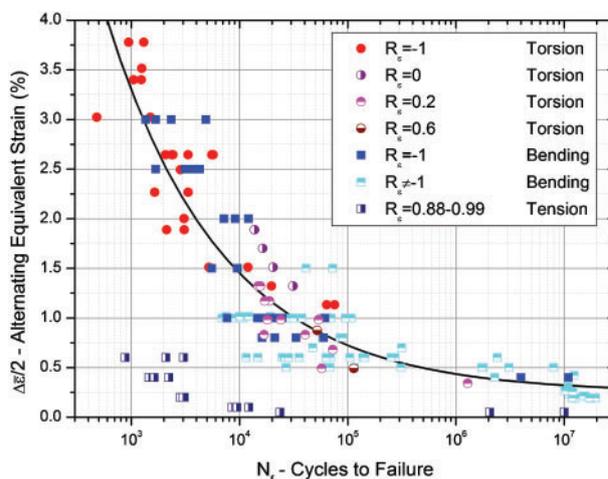


- 23 Stent pulsatile fracture data are overlaid on the diamond fatigue data from Fig. 20. The open circles represent test conditions for SMART stents that survived 10^7 pulsatile cycles. The closed circles are those conditions that led to fracture of the SMART stent at $<10^7$ pulsatile fatigue cycles. Note that the stent fatigue testing is consistent with the diamond data, whereby fracture tends to occur above the worst case 0.4% strain amplitude (dotted line) for a range of mean strains. The shaded box represents mean and amplitude strains for typical endovascular oversizing and pulsatile fatigue conditions. Reprinted from Pelton *et al.*⁴⁶ with permission from Elsevier

between ~ 0.2 and 0.45% that are closer to physiological conditions (approximately shown with a shaded grey box) and are consistent with the worst case 0.4% fatigue strain limit line. The results shown in Fig. 23 demonstrate the incorporation of the diamond subcomponent testing analysis as a useful and accurate indicator of Nitinol stent lifetimes for radial dilation fatigue conditions.

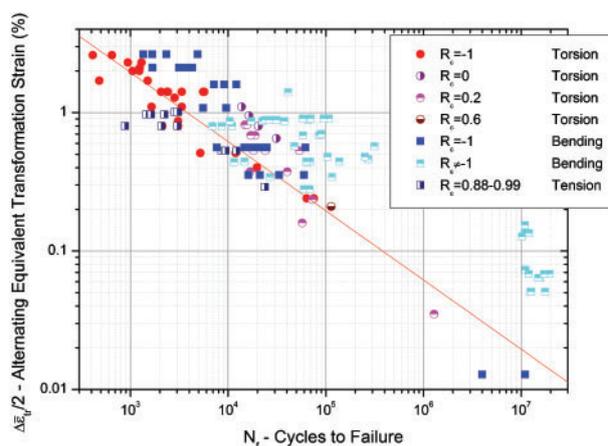
Multiaxial or mixed mode conditions

The above examples of strain controlled fatigue for Nitinol medical devices were carried out under conditions of uniaxial or bending cycles. However, to better understand the complex *in vivo* performance of Nitinol medical devices, a more complete multiaxial model must be developed. Torsion of thin walled tubes provides a convenient method to characterise the effects of shear stress and strain. A recent investigation cycled Nitinol tubing in torsion at various mean and alternating strains and compared the results with equivalent fatigue data collected under uniaxial tensile/bending conditions.⁵⁸ In order to compare the torsional hysteresis and fatigue data with previous results obtained under axial tension and bending, the shear strain values were converted to equivalent strains. The results were expressed as the equivalent true (Cauchy) stress $\bar{\sigma}$ and the equivalent referential (Lagrangian) strain $\bar{\varepsilon}$. The data were converted by using $\bar{\sigma} = (\sigma_t^2 + 3\sigma_s^2)^{1/2}$ and $\bar{\varepsilon} = (\varepsilon_t^2 + 4/3\varepsilon_s^2)^{1/2}$, where σ_t and ε_t are the tensile stress and strain respectively and σ_s and ε_s are the shear stress and strain respectively. For the deformation range of interest in that study, i.e. mean strains $<5\%$, the transformation strains ε_{tr} are defined as the total strain minus the (purely) elastic strain. Replotting the diamond and 'strut' fatigue data from Fig. 23 along with the Nitinol tube torsion data for R_e values of -1 , 1 ,



- 24 Combined ε - N diagram for the multiaxial fatigue of superelastic Nitinol tubes plotted in terms of the number of cycles showing normalisation of multiple datasets of torsion, bending (tension/compression) and tension-tension fatigue results at multiple R_e ratios. This formulation does not appear to work so well for mean strains on the order of 9% ($R_e \approx 0.88-0.99$). Reprinted from Runciman *et al.*⁵⁸ with permission from Elsevier

0.2 and 0.6 on a single alternating equivalent strain versus number of cycles, strain-life diagram displays a remarkable normalisation of superelastic Nitinol fatigue data, as shown in Fig. 24. However, the tension-tension data with high mean strain data do not follow the curve. An alternative and improved approach that can accommodate the high mean strain results is to incorporate a modified Coffin-Manson type equation for the multiaxial fatigue of Nitinol using the number of cycles and equivalent transformation strain amplitude. The equivalent transformation strain $\bar{\varepsilon}_{tr}$ is defined as the total equivalent strain $\bar{\varepsilon}$ minus the elastic equivalent strain $\bar{\varepsilon}_{el}$, where $\bar{\varepsilon} = \bar{\varepsilon}_{el} + \bar{\varepsilon}_{tr}$; the transformation strain amplitude is therefore $\Delta\bar{\varepsilon}_{tr}/2$. Applying this notion to the torsion, bending and



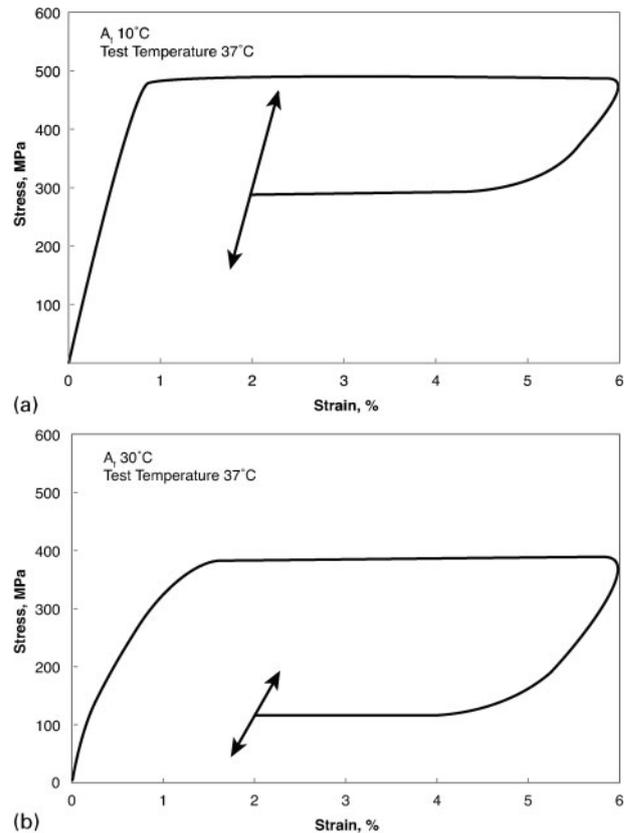
- 25 A modified Coffin-Manson relationship for the multiaxial fatigue ($-1 < R_e < 0.99$) of superelastic Nitinol involving the alternating equivalent transformation strain $\Delta\bar{\varepsilon}_{tr}$ as a function of the number of cycles to failure, N_f , where $\Delta\bar{\varepsilon}_{tr}/2 = 61.7N_f^{-0.5}$. Note how this approach normalises the ε - N data even for the highest mean strain results out to 9%. Reprinted from Runciman *et al.*⁵⁸ with permission from Elsevier

tension–tension fatigue data of superelastic Nitinol in Fig. 24 over a wide range of strain ratios from R_ϵ values of -1 out to 0.9 revealed a ‘universal’ fit (Fig. 25), which can be expressed in terms of the following relationship: $\Delta\bar{\epsilon}_{tr}/2 = 61.7N_f^{-0.5}$. Accordingly, this phenomenological approach provides a means to normalise fatigue conditions from a number of deformation modes for thermomechanically processed Nitinol used to manufacture medical devices.

Energy considerations

Mechanistically, it is not known if the fatigue behaviour observed in superelastic Nitinol is due to microstructural effects of martensite (i.e. formation of stabilised martensite during cycling with the incorporation of plasticity) or rather to the lower moduli associated with the stress induced transformation regime (i.e. decreased hysteresis energy).⁴⁶ In this respect, however, there is merit to consider the details of the cyclic hysteresis loops for predictive purposes. This approach is incorporated, for example, in thermodynamic analyses of thermal and mechanical fatigue of Nitinol where changes in transformation temperature and upper and lower transformation stresses are used to model the elastic and irreversible energy as a function of fatigue cycles.^{59–61} The change in elastic free energy relationship from these analyses was expressed as: $\Delta E_e \propto \int \sigma_{ij}^t \epsilon_{ij}^t dV/V - \int (\sigma_{ij} + \sigma_{ij}^d) \epsilon_{ij}^t dV/V$, where σ , σ_d and σ_{tr} are the applied stress, the internal elastic stress field due to the presence of defects and the internal stress field associated with the transformation respectively. For tension–tension superelastic cycling, these models are consistent with decreasing elastic energy with cycles due to the increase in internal elastic stress field in the direction of the applied stress resulting from creation and rearrangement of dislocations and deformed martensite. The area within the hysteresis loops from superelastic cycling is a measure of work and therefore can be modelled in an analogous manner to account for damage accumulation due to the effects of mean strain (stress) and strain (stress) amplitude. Smith–Watson–Topper (SWT)⁶² developed a phenomenological method to incorporate mean stress for the analysis of fatigue data. They hypothesised a generalised stress–strain function to predict fatigue life and assumed that the function should be applicable to crack initiation and at least the early part of crack propagation**. This function is based on determining the stress range $\Delta\sigma$, strain range $\Delta\epsilon$ and modulus of elasticity E . One version of this relationship for strain controlled tests can be stated as $P^{SWT} = (\sigma_{max} \Delta\epsilon E)^{0.5}$; other forms of the relationship are applicable for stress controlled tests under conditions of fully reversed and with application of mean strain (stress).⁶² From this simple relationship, it can be seen that the fatigue parameter P^{SWT} is minimised with minimal values of stress, strain and modulus. A modified SWT analysis was used to analyse Nitinol fatigue data thermal cycles under conditions of a constant applied force.⁶⁷ The modulus for Nitinol is extremely temperature dependent²⁴ and changes with the value of mean strain due to the relative amount of austenite and martensite. Nevertheless, a stabilised

**As we will discuss in the subsequent section, Nitinol has a very low fatigue crack threshold.^{43,63–66} Therefore, once a crack initiates, it takes very little additional energy to propagate the crack. Consequently, the SWT analysis is suitable for total life fatigue of Nitinol.



26 Schematic stress–strain curves with strain controlled fatigue conditions of 2% mean strain and 0.25% strain amplitude for Nitinol tested at 37°C for A_f values of a 10°C and b 30°C. Note that the effective modulus for each condition is comparable with the initial loading modulus. Therefore, although the fatigue strain conditions are identical, the stress ranges are dramatically different. The total energy per cycle, represented by the arrow length, reduces with increasing A_f (decreasing ΔT) with a corresponding reduction in microstructural damage accumulation and increase in fatigue life

modulus can be used for these analyses. For the case of typical *in vivo* conditions of mean strain and strain amplitude, it was shown in Figs. 13 and 15 that the cyclic stress–strain curves adopt a modulus comparable with the initial loading modulus (especially with low values of mean strain and strain amplitude); this is also observed in Fig. 7 for strain controlled conditions of thermal martensite, per MM1. We can, therefore, schematically represent these stabilised fatigue hysteresis loops with a straight line as illustrated in Fig. 26 for the conditions of 2% mean strain and 0.25% strain amplitude with A_f values of 10 and 30°C and tested at 37°C (i.e. tested in a superelastic condition). This figure summarises the salient features observed in several investigations. First, the loading and unloading plateau stresses decrease with increasing A_f temperature (decreasing ΔT). Second, the cyclic modulus decreases with increasing A_f . The modulus depends upon the extent of prestrain, mean strain (stress) and strain (stress) amplitude; a first order approximation of the modulus can be estimated with an austenite–martensite rule of mixtures. Furthermore, these parameters depend on the processing conditions, as shown in Figs. 13 and 15. The stress–strain characteristics shown in Fig. 26b result

in predictions of longer fatigue lives with increasing A_f , whereby a work integral or SWT function is minimised with reduced damage accumulation as demonstrated with the data in Figs. 8–10. Clearly, these approaches are conceptual and require additional vetting to be used quantitatively to predict fatigue life in Nitinol. It will be especially important for detailed mechanical and microstructural analyses to be carried out in the future for varying conditions of mean strain (stress), strain (stress) amplitude and transformation temperature to add additional insight into Nitinol fatigue. Examples are shown in Figs. 13–15 to illustrate microstructural differences by TEM. Moreover, it was demonstrated recently that synchrotron micro-X-ray diffraction could be used to differentiate elastic and transformational strains in deformed Nitinol.^{31,68} This characterisation method is capable of $\sim 1 \mu\text{m}$ resolution, which could provide a high resolution technique to monitor mechanically fatigued Nitinol.

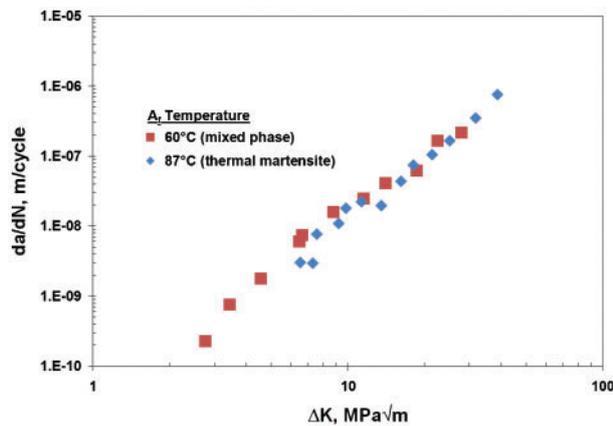
Furthermore, none of these changes in material response have yet to be incorporated in traditional FEA stress computations for Nitinol components, although Mehta *et al.* addressed some of these discrepancies by direct measurement of the transformation using *in situ* synchrotron X-ray diffraction analysis.⁶⁸

Summary of trends with stress/strain–life data

1. Superelastic Nitinol provides longer lives and greater fatigue limits than thermal martensite under stress controlled conditions. These observations are attributed to the difference in plateau strength between austenite and martensite at the test temperature. The stress fatigue limit is $\sim 80\%$ of the respective stress plateau (superelastic) or detwinning plateau (thermal martensite). Furthermore, superelastic conditions produce only a few of the possible 24 variants of martensite that are aligned to optimise internal strain distributions and minimise dislocation formation.
2. Repeated cycles under stress controlled conditions lead to deformational changes, i.e. an increase in peak strain at constant stress amplitude. The magnitude of these deformational changes depends on processing and test conditions.
3. The fatigue life of thermal martensite is superior to superelastic Nitinol under strain controlled conditions. This is attributed to the fact that for a given strain amplitude, the stress required to deform thermal martensite is lower than those required to transform austenite to martensite under superelastic conditions.
4. Repeated cycles under strain controlled conditions lead to cyclic hardening, i.e. an increase in stress at constant strain amplitude. This hardening is due to the effects of an increase in dislocation density with increasing cycles that tends to reduce the amount of material undergoing cyclic phase transformation.
5. Strain controlled fatigue testing of medical device subcomponents can be described by a non-traditional constant life diagram. Fatigue life for superelastic Nitinol is driven by strain amplitude rather than mean strain or peak strain. Furthermore, fatigue life is increased with increasing mean strain up to the end of the phase transformation plateau.
6. Strain–life approaches based on the alternating equivalent strain, or better still the alternating equivalent transformational strain incorporated into a modified Coffin–Manson formulation, provide a means to analyse fatigue data for Nitinol tested under multiaxial or mixed mode loading conditions.
7. Thermomechanically processed Nitinol is more stable than fully annealed conditions in which dislocations accumulate more rapidly in the large grained annealed material, which results in shorter lives.

Damage tolerant fatigue

Compared with the total life approach to fatigue of Nitinol, relatively few studies have been conducted to evaluate the damage tolerant properties, i.e. the fatigue crack growth and fracture toughness behaviour. One reason for this is that currently the Nitinol industry is dominated by commercial products with geometrically small feature sizes where the emphasis is placed on preventing fatigue crack nucleation rather than controlling crack growth. Not only is it difficult to experimentally measure crack growth rates accurately in such physically small components, but once a macroscopic crack nucleates and starts growing, failure generally ensues shortly afterwards as is shown in the body of work presented herein. For example, a stent with a typical strut width of $\sim 0.2 \text{ mm}$ cycled at 1 Hz (heart rate) may require only hours or days for a crack to fully propagate through the wall once nucleated. Indeed, in those cases, it is important to screen, and reject, devices for surface flaws that are larger than the critical crack length for such propagation, which in stents is typically about $15\text{--}50 \mu\text{m}$.⁴³ Provided that those geometrically small devices do not have flaw sizes larger than this critical crack length, then the total life approach to fatigue evaluation is more appropriate. Furthermore, the Food and Drug Administration, which regulates these devices, requires testing of complete commercial devices, rather than solely of ‘coupon specimens’, such that there is a direct correlation between total life ‘bench top’ fatigue lifetimes and in service device performance. As discussed below, the damage tolerant approach to fatigue is most appropriate when the product is in a stress controlled environment and its geometric size is large enough to sustain stable crack growth for many thousands of cycles and over a long enough time period to retain functionality of the device over its lifetime despite the presence of a growing crack, or when devices cannot be adequately examined and rejected for surface flaws exceeding the critical crack length. In today’s commercial landscape, there are few Nitinol products that meet such a description. One such example is the concrete reinforcing bars for rapid rebuilding using the shape memory effect following structural damage to bridges. However, damage tolerant analyses are useful as a compliment to a total life approach. Specifically, whereas stress/strain life data should be employed to provide the primary basis for life prediction in most medical devices, damage tolerant analyses can be used as secondary life prediction methodology that can additionally assess, in the quantitative way, the effect of the presence of defects in terms of reduced performance/lifetimes.⁶⁹ Moreover, the commercial landscape may



27 Fatigue crack growth da/dN data as a function of the stress intensity range ΔK for stress relieved Nitinol alloys where the stable phase was controlled by varying the chemical composition. The crack growth rates were virtually identical between the two microstructures tested at room temperature with $R < 0.1$; after MM1¹³

change in the future, and require the implementation of damage tolerant techniques as the primary life prediction strategy. Indeed, the orthopaedics community is beginning to adopt Nitinol for *in vivo* skeletal stabilising devices (in a stress controlled environment) that have geometries on the order of several centimetres, where the damage tolerant analysis is more relevant.

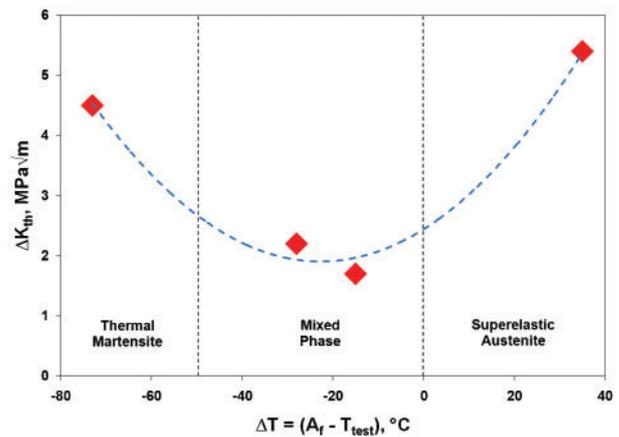
Phase dependent fatigue crack growth

MM1¹³ were the earliest researchers to conduct fracture mechanics based fatigue crack growth measurements on Nitinol. Their material was upset forged at 800°C from an initial 50 mm diameter ingot into an 18 mm thick plate, followed by a 1 h anneal at 950°C. This treatment results in a distinctly different large grain structure than is found in most modern commercial applications that use heavily cold worked and subsequent stress relieved Nitinol (as discussed in the previous section). In that study, two different Nitinol compositions were examined at room temperature, namely a thermal martensite alloy ($M_s = 47^\circ\text{C}$, with $A_f \approx 87^\circ\text{C}$), and a mixed phase material ($M_s = 20^\circ\text{C}$, with $A_f \approx 60^\circ\text{C}$). Despite the radically different microstructures, they found virtually no difference in the crack growth behaviour of the two materials; Paris law^{††} exponents were $m \approx 3.5$, with a threshold stress intensity value of $\Delta K_{th} \approx 2.4 \text{ MPa m}^{1/2}$ ^{‡‡} (Fig. 27).

It was a decade later until further studies on fatigue crack growth behaviour in Nitinol were reported; specifically two studies were presented at an MRS International Meeting on ‘Advanced materials’ in Japan in 1988. The first of these, by Dauskardt *et al.*,⁷⁰ varied the Nitinol chemical composition in order to obtain various microstructures at room temperature. Whereas those authors identified the material as ‘stable austenite’,

^{††} The Paris power law provides a means to relate the fatigue crack growth rates da/dN to the stress intensity range ΔK . In its simplest form, it can be expressed as: $da/dN = C\Delta K^m$, where C and m are experimentally determined scaling constants. The latter expression is most applicable to intermediate growth rate behaviour typically between $\sim 10^{-6}$ and 10^{-9} m/cycle.

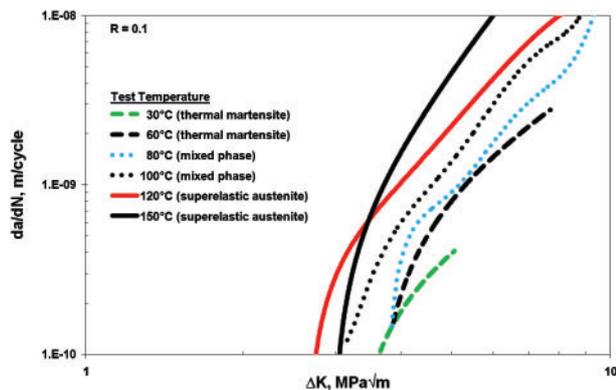
^{‡‡} The fatigue threshold stress intensity ΔK_{th} represents the value of the stress intensity range below which (large) crack growth is essentially dormant.



28 Dauskardt *et al.*⁷⁰ fatigue crack growth threshold stress intensity values in Nitinol plotted versus ΔT (defined above). The data show that mixed phase material (those with ΔT just below zero) has a lower fatigue threshold than both the thermal martensite ($\Delta T = -73^\circ\text{C}$) and the superelastic austenite ($\Delta T = +35^\circ\text{C}$)

‘irreversible transforming superelastic austenite’, ‘reversible transforming superelastic austenite’ and ‘stable martensite’, evaluation of the A_f and test temperatures revealed that the authors studied superelastic austenite, mixed phase material and thermal martensite. All microstructures demonstrated similar Paris law growth exponents (2.7–3.5), but the ΔK_{th} threshold values were distinctly different. Specifically, the superelastic austenite and thermal martensite microstructures had much larger (by a factor of ~ 2) fatigue thresholds, and hence superior fatigue crack growth resistance, than the mixed phase materials (Fig. 28). Because of the confusion of the microstructures, those authors incorrectly concluded that superelastic austenite has the lowest fatigue threshold value and ascribed threshold differences to a negative volumetric change associated with the phase transformation from austenite to martensite that they termed ‘negative dilation’. They claimed further that, due to the constraint of surrounding untransformed materials remote from the growing crack, negative dilation in superelastic austenite would effectively raise the local stress intensity immediately ahead of the propagating crack (negative crack tip shielding). In other words, the decrease in volume of the transformed material immediately ahead of a growing crack induces an elevated tensile strain field in the surrounding untransformed grains due to the required need to compensate (by tensile elastic loading) for the shrinkage of the transformed material. While this hypothesis is, indeed, a valid scientific argument, the authors mistakenly identified the superelastic austenite as a stable austenite. Therefore, their argument that a transforming microstructure offers poorer threshold stress intensities is incorrect since the superelastic austenite material demonstrated the largest threshold value. Instead, we opine that the cyclic fatigue in a mixed phase microstructure creates a complex transformation structure (based on the energetic competition between mechanically and thermally formed martensitic variants previously discussed) that causes easier fatigue crack initiation and consequently lower threshold stress intensity values.

The second study presented at the 1988 MRS meeting was from Miyazaki *et al.*⁷¹ who investigated fatigue

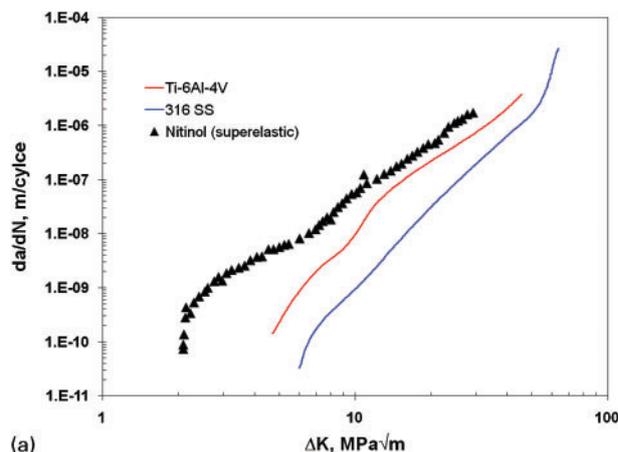


29 Fatigue crack growth data for various Nitinol microstructures controlled by modifying the test temperature. Whereas the stable crack growth exponents were largely equivalent, the threshold stress intensity for the superelastic austenite was lower than the thermal martensite; after Holtz *et al.*⁷²

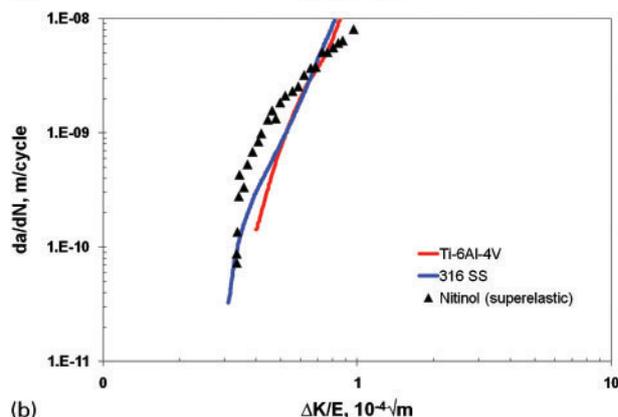
crack growth behaviour in Nitinol plate evaluated in the Paris regime (i.e. at intermediate growth rates) as a function of temperature, chemical composition and heat treatment. These authors noted the importance of modifying such variables to obtain the various microstructural phases, but also recognised that each of these changes can introduce new biases into the interpretation of the data. Unfortunately, the results presented by Miyazaki *et al.* show Paris law exponents m that are much higher, and constants C that are much lower than any other studies in the literature on the fatigue of Nitinol. Indeed, the Paris law exponents were some 3–4 times higher, and Paris constants 5–10 orders of magnitude lower than any other values reported by the rest of the Nitinol fatigue testing community. The wild discrepancy between these data and the rest of the body of Nitinol literature leads us to conclude that the values that were calculated are erroneous, and accordingly we are compelled to ignore these data. Their data are tabulated for completeness later, but are excluded from the trend studies.

A later fatigue study, by Holtz *et al.*,⁷² involved holding the chemical composition constant and varying the test temperature (from 30 to 150°C) to achieve specific phases in a binary $\text{Ti}_{50.1}\text{Ni}_{49.9}$ (at-%) alloy, heat treated for 1 h at 750°C with an $A_f=120^\circ\text{C}$. This work involved a very complete test matrix of two phases (thermal martensite and superelastic austenite) and four load ratios of $R=0.1-0.9$. As expected, the fatigue threshold ΔK_{th} values decreased with increasing load ratio (Fig. 29). Similar to the conclusions of Dauskardt *et al.*,⁷⁰ they found that thermal martensite microstructures required the highest stress intensities for crack propagation, and hence were the most resistant to fatigue crack growth.

McKelvey and Ritchie⁶³ investigated the fatigue crack growth rates of binary $\text{Ni}_{50}\text{Ti}_{50}$ (at-%) alloy in the superelastic condition in thick section (9 mm thick) round compact tension samples (machined from 41.3 mm diameter round bar) tested in 37°C Hank's balanced salt solution at 10 Hz frequency with $R=0.1$. They found that, compared with other biomedical implant materials, superelastic Nitinol had the lowest fatigue crack growth resistance in terms of the highest growth rates (at a given



(a)

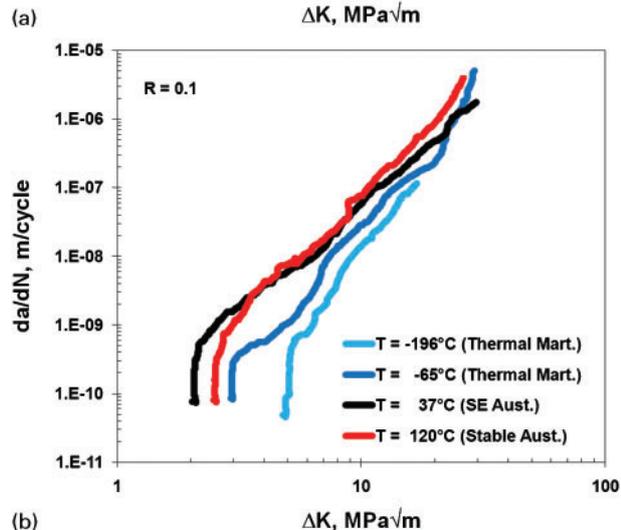
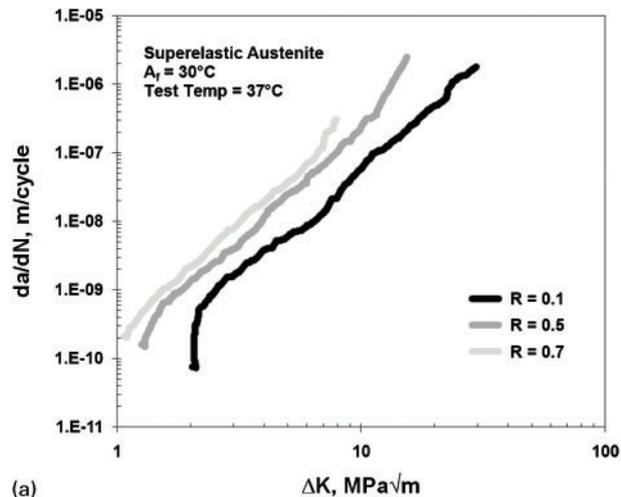


(b)

30 Comparison of fatigue crack growth rates for various biomedical metallic alloys showing a distinct inferiority of Nitinol when evaluated by stress intensity (after McKelvey and Ritchie⁶³), but b equivalency threshold values when evaluated by strain intensity (i.e. when data in Fig. 28a are normalised by the elastic modulus) in the nominal linear-elastic range. Elastic modulus for the titanium alloy (Ti-6Al-4V), austenitic stainless steel (316SS) and Nitinol were taken as 117, 193 and 62 GPa respectively

ΔK) and the lowest threshold values ($\Delta K_{th} \approx 2 \text{ MPa m}^{1/2}$) (see Fig. 30a). However, if one follows the approach of Ohta *et al.*⁷³ and normalises the fatigue crack growth rates by the modulus of the material, thereby evaluating the strain intensity range (rather than the more traditional stress intensity range), the Nitinol data from McKelvey *et al.* can be shown to approximately the same as comparative biomedical metallic implant materials such as Ti-6Al-4V and austenitic stainless steels (Fig. 30b). This trend suggests that Nitinol displays equivalent fatigue crack growth properties to other biomedical metallic implant materials under strain controlled boundary conditions, especially in the slow growth regions near the threshold values.

McKelvey and Ritchie later compared the relative fatigue resistance of the various different structural phases in Nitinol for this thick section material, and further investigated the role of R ratio (from 0.1 to 0.7) on fatigue behaviour.⁶⁴ Using temperature (rather than composition) to change the structural phase in a single $\text{Ni}_{50}\text{Ti}_{50}$ (at-%) alloy (with $A_f \approx 30^\circ\text{C}$), they compared the fatigue crack propagation behaviour over a wide range of growth rates (from 10^{-11} to 10^{-5} m/cycle) in stable austenite (at 120°C), superelastic austenite (at



31 Variation in fatigue crack propagation rates da/dN as a function of the stress intensity range ΔK in an A_f of $\sim 30^\circ\text{C}$ Nitinol tested in thick section (9 mm thick compact tension specimens) at various temperatures at 10 Hz frequency. The results show *a* the effect of load ratio ($R=0.1-0.7$) in the superelastic austenite microstructure at 37°C and *b* the relative fatigue crack propagation resistance of the stable austenite, superelastic austenite and thermal martensite microstructures; after McKelvey and Ritchie⁶⁴

37°C) and thermal martensite (at -65 and -196°C). Their observations on the effect of load ratio, shown in Fig. 31*a* for superelastic austenite at 37°C , demonstrated how the fatigue thresholds decrease and growth rates increase (particularly at high and near threshold levels) with increasing R . The best fatigue crack growth resistance, in terms of lowest growth rates and highest thresholds, was found in thermal martensite. Furthermore, the stable austenite structure showed slightly better properties than superelastic austenite (Fig. 31*b*). Although these comparisons have to be made at different temperatures, which may produce some degree of bias in the conclusions, the observation that thermal martensitic has superior fatigue crack propagation resistant compared with superelastic austenite, is entirely consistent with the earlier studies of Holtz *et al.*⁷² described above.

Effects of inclusions

The specific effect of inclusions on fatigue crack propagation in Nitinol was examined by Vaidyanathan

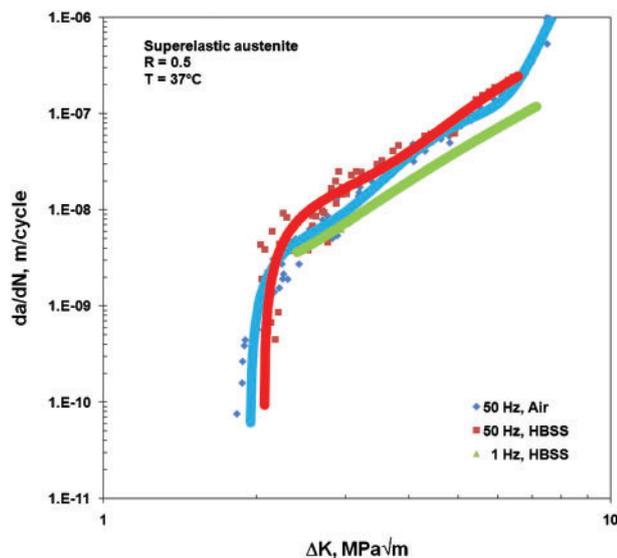
*et al.*⁷⁴ who studied the role of increasing the density of titanium carbides to 10–20 vol.-% in a superelastic austenitic alloy. They concluded that increasing the inclusion content increased the fatigue threshold stress intensity. We opine, however, that with an inclusion content this high, the depletion of titanium from the bulk matrix would inherently change the material from binary superelastic Nitinol to an alloy with a Ni rich stoichiometry that cannot be compared directly with commercially available Nitinol material. Furthermore, modern commercially produced Nitinol materials are designed to contain low inclusion (both oxides and carbides) contents that would never approach the 10–20% range studied in this work. It should be noted, though, that the ‘inclusion free’ superelastic Nitinol that they did examine in this work had a crack growth exponent m similar to other literature values, but that the threshold stress intensity was 30–110% higher than other reported values tested at $R=0.1$.

The presence of near surface inclusions often provides preferential nucleation sites for fatigue fracture both in laboratory experiments and in commercial devices; see the later section on ‘Fractography’ for examples. However, one should not mistakenly conclude that the presence of these inclusions necessarily promotes poor fatigue resistance. Indeed, although inclusions certainly cause localised stress discontinuities, the size of inclusions in contemporary Nitinol products is typically smaller than the 15–50 μm critical flaw size for Nitinol⁴³ and therefore may have limited, if any, effect on the fatigue crack growth properties. This conclusion is supported by comparing the total life rotary bending fatigue work of Sheriff *et al.*³⁶ in Nitinol manufactured with ASTM acceptable inclusion levels to Morgan *et al.*'s³⁸ extra low inclusion Nitinol and Reinoehl *et al.*'s³⁵ work with black oxide wire, which show statistically similar low and high cycle fatigue behaviour under fully reversed alternating strain conditions (see Fig. 10 for graphical comparison)⁵⁵.

Role of environment

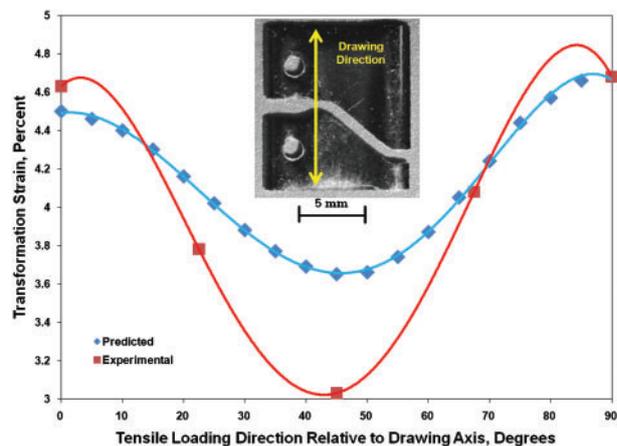
In a temperature dependent phase transforming material such as Nitinol, the experimental set-up can have just as much influence on the fatigue crack growth behaviour as the chemical composition and thermomechanical processing. Indeed, even small changes in the test temperature can have a profound effect on Nitinol's mechanical properties (see e.g. Fig. 1), which is why it is essential to characterise the transformation temperatures. Despite this sensitivity, McKelvey and Ritchie⁶³ and later Robertson and Ritchie⁶⁶ in a follow-up study to the earlier work of Stankiewicz *et al.*⁶⁵ came to an important conclusion that adiabatic heating due to the superelastic phase transformation had little influence on the crack growth behaviour. These cumulative works showed no differences in samples tested in air (with low thermal conductivity thereby allowing the possibility of adiabatic heating) and a simulated body fluid (with high thermal conductivity which would limit any adiabatic heating ahead of the growing crack tip). Furthermore,

⁵⁵ Whereas Sheriff's, Reinoehl's and Morgan's experimental testing in physically small specimens seems to circumstantially confirm the argument that inclusions are subcritical in dimension, there have been no experimental research performed to directly quantify the fatigue nucleation effect of inclusions whose dimensions represent a significant portion of the total cross-sectional area. This topic is discussed further in the small versus large crack section.



- 32 Fatigue crack growth curves for electropolished superelastic austenite Nitinol tested at 37°C and $R=0.5$ in air and Hanks' balanced saline solution (HBSS) at accelerated frequencies (50 Hz), and in HBSS at physiologically relevant frequency (1 Hz). The curves show no statistically significant difference, thereby suggesting that adiabatic heating due to the phase transformation has no effect on the crack growth rates and that corrosion assisted fatigue is not active in Nitinol at these frequencies. After Robertson and Ritchie⁶⁶

McKelvey and Ritchie,⁶³ Filip *et al.*⁷⁵ and Robertson and Ritchie⁶⁶ all reached the same conclusion that the testing of Nitinol at temperatures of 37–40°C in a simulated body solution, as compared with air, does not adversely influence the crack growth behaviour of electropolished specimens, indicating that corrosion assisted fatigue is not of concern under those specific circumstances. This is an important issue as most researchers, for the sake of expediency, employ cyclic testing frequencies well above expected in service conditions, which would have the potential of suppressing any such corrosion assisted fatigue (if it were active). This specific issue was resolved by Robertson and Ritchie⁶⁶ who fatigue tested identically processed superelastic austenite specimens at 1 Hz and at 50 Hz in 37°C simulated body fluid and found essentially identical fatigue crack growth behaviour,⁶⁶ thereby strengthening the conclusion that electropolished Nitinol is not susceptible to corrosion assisted fatigue in a simulated human blood solution at physiological temperatures over this frequency range (Fig. 32). Although these conclusions are encouraging to researchers who have utilised elevated frequencies for rapid data collection, we must caution that these results may not be applicable to all instances of Nitinol fatigue. For example, gastrointestinal Nitinol devices are subjected physiologically to strongly acidic aqueous environments where corrosion may play a very important role in the rate of fatigue crack growth; consequently, it is important to select the appropriate pH simulated solution and testing frequency to allow for the possibility of corrosion assisted fatigue to occur. In the previous section, we showed experimental evidence showing that the stable fatigue crack growth slopes are

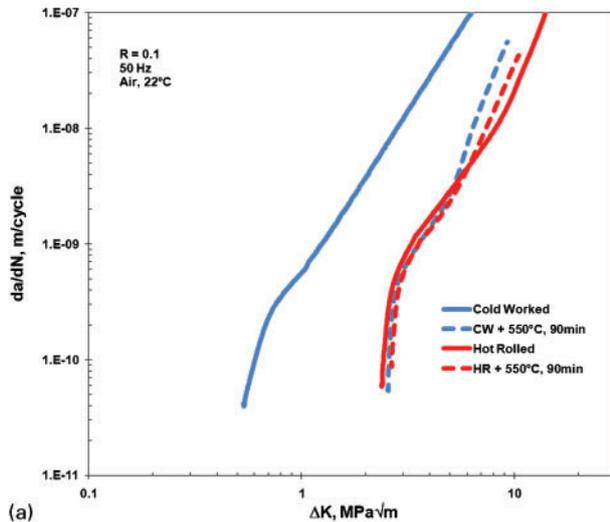


- 33 Orientation dependent transformational strains predicted and measured experimentally by Robertson *et al.* to explain the unexpected preferential fatigue crack growth behaviour in Nitinol compact tension specimens (inset photo); after Robertson and Ritchie⁶⁶

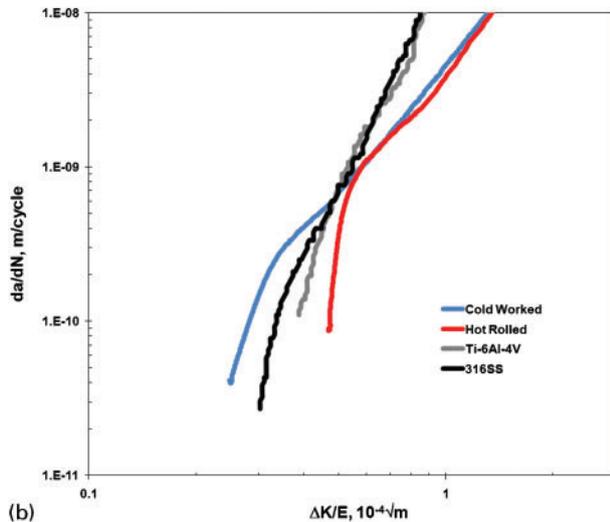
independent of the load ratio. Therefore, although these tests were conducted only at $R=0.5$, we believe that electropolished Nitinol's resistance to corrosion assisted fatigue in this simulated body fluid environment spans to all load ratios.

Texture dependent fatigue crack growth

The role of crystallographic texture on the mechanical properties of superelastic Nitinol is well known and manifests as widely varying strains required to stress induce the martensitic phase depending on the material texture.^{20,21,76} This texture develops from the casting conditions as well as from the extensive thermomechanical processing to the final product configuration. Not only is texture responsible for the asymmetry in mechanical properties between tension and compression,²⁹ but it varies significantly between different product forms⁷⁷ in Nitinol. This is important to realise as the texture in Nitinol is practically impossible to modify in the final form, even using aggressive annealing and/or thermomechanical treatments. It is, therefore, expected that the fatigue crack growth properties of superelastic Nitinol would also be affected by texture. Indeed, reports in the literature demonstrate this dominance of the role of texture on the crack growth path in thin walled tubing that has undergone several thousands of percentage of reduction from its original as cast ingot form. Specifically, Robertson and Ritchie⁶⁶ showed that the preferred (i.e. lowest crystallographic energetic barrier) crack growth direction in thin walled Nitinol tube used for endovascular stents was 45° to the tube drawing direction (Fig. 33). Surprisingly, cracks grew at 45° to the pure mode I direction in samples cut with the prenotch oriented parallel or perpendicular to the drawing direction, demonstrating the marked dominance of texture on fatigue crack growth in superelastic austenite. In another study, Robertson *et al.*⁷⁸ characterised the texture and strain distribution immediately ahead of a growing fatigue crack in superelastic austenite using synchrotron X-ray diffraction techniques. These authors demonstrated that the strain distribution and transformation did not follow the expected peanut shaped front that is predicted from linear-elastic fracture mechanics models. Instead, grains with certain orientations could suppress the transformation even when surrounding grains had already undergone the



(a)



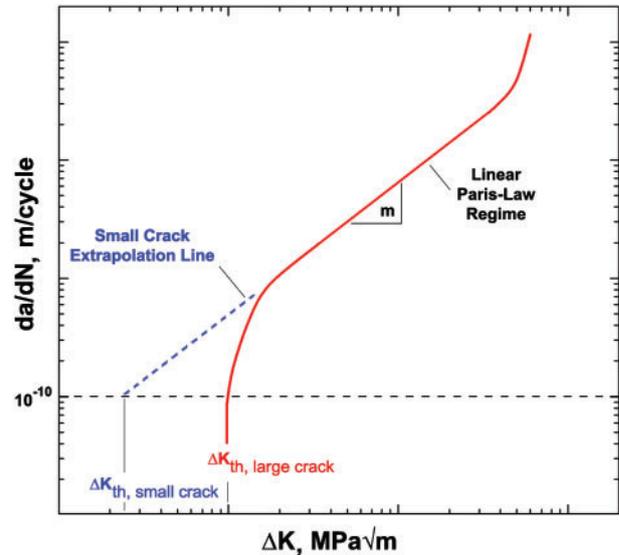
(b)

- 34 *a* fatigue crack growth behaviour of wrought Nitinol bars showing a trend of reduced threshold stress intensity values in materials with higher amounts of dislocation density, and *b* threshold data normalised by material modulus to compare the strain intensity values of wrought Nitinol with other biomedical metals; after Gall *et al.*⁷⁹ Cold worked Nitinol was stable austenite, cold worked plus heat treatment was mixed phase and both of the hot rolled conditions were in the superelastic austenite condition. Elastic modulus values for the Ti-6Al-4V, 316SS, cold worked wrought Nitinol and hot rolled wrought Nitinol were 117, 193, 21 and 51 GPa respectively

transformation. These two important observations highlight the necessity to evaluate the fatigue crack growth properties in user's specific product form rather than assuming that the reported literature values are appropriate for their product.

Effect of hot and cold work

The fatigue crack growth data presented by Gall *et al.*⁷⁹ again highlight the criticality of reporting the processing parameters. They compare hot rolled bar Ti_{49.1}Ni_{50.9} samples to bars that had been both hot rolled and then subsequently cold drawn 30%. Four samples with various aging treatments from each configuration were examined: as formed, 300°C for 90 min, 350°C for 90 min and 550°C for 90 min. As expected, the cold drawn material with its high inherent dislocation density



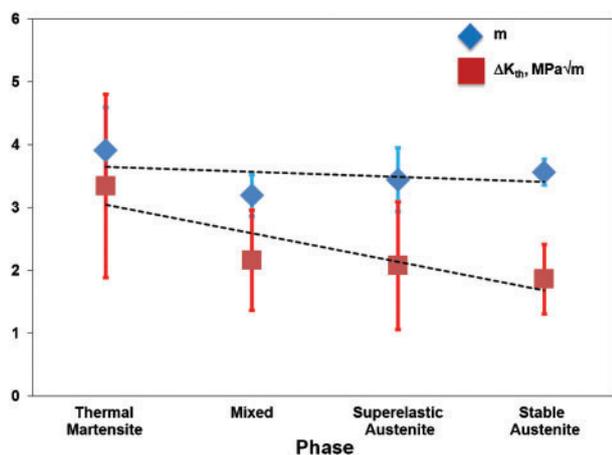
- 35 Diagram showing a means to estimate approximate small crack fatigue threshold values by extrapolating the corresponding large crack Paris regime data down to 10^{-10} m/cycle; after Robertson and Ritchie⁴³

required the least amount of initiation energy and had the lowest threshold stress intensity. What was surprising though was the extremely low magnitude of their measured threshold values. Indeed, in the as formed cold drawn bar and the mildly heat treated samples (300 and 350°C for 90 min, both temperatures well below the onset of recrystallisation at $\sim 450^\circ\text{C}$), the ΔK_{th} values were $< 1 \text{ MPa m}^{1/2}$, which is extremely low for most other engineering metals and even lower than for many ceramics (Fig. 34a). The effect was deemed to be associated with residual stresses resulting from prior cold work. However, the threshold stress intensity increased to 'normal' levels after the cold worked material was aged at temperatures high enough to promote recovery and recrystallisation (550°C for 90 min). This observation demonstrates the importance of conducting an aging anneal above the recrystallisation temperature in cold drawn Nitinol to reduce the high dislocation density, thereby imparting a mechanism of fatigue crack nucleation resistance by allowing the cyclic build-up of dislocations. Unlike the McKelvey and Ritchie⁶³ data, the threshold stress intensity region of the crack growth curves did not overlap with other biomedical metallic implant materials when the data were normalised by the material modulus (Fig. 34b).

Small versus large cracks

The role of small versus large cracks^{¶¶} has been notably absent from the discussion of Nitinol fatigue. Robertson and Ritchie⁴³ estimated, by extrapolating the linear Paris regime curve down to 10^{-10} m/cycle (Fig. 35), that the fatigue crack growth threshold values for physically small cracks were likely some 40–60% lower than for

¶¶ Cracks are considered to be small when they are: microstructurally small, i.e. comparable in size with the scale of the characteristic microstructure (a continuum limitation), mechanically small, i.e. comparable with the scale of local plasticity (a linear-elastic limitation), or functionally small, i.e. comparable with the size of the local crack tip shielding zones (e.g. from crack closure or crack bridging) that can develop in the wake of the crack tip (a crack tip similitude limitation). Compared with the behaviour of corresponding large cracks at equivalent stress intensity range, small cracks generally display elevated growth rates and lower fatigue threshold values.⁸⁰



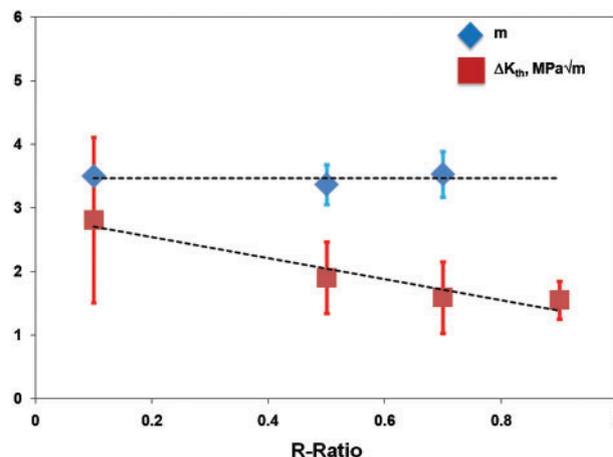
36 Crack growth parameters (Paris exponent m and fatigue threshold ΔK_{th}) for Nitinol as a function of the structural phase. Data are presented as mean \pm one standard deviation. Exact transformation times and temperatures are shown in Table 3

correspondingly large cracks (Table 2). However, there have been no reported experimental tests that actually have measured small crack growth and small crack thresholds in Nitinol alloys. Given the current commercial practice of using Nitinol for physically small devices, e.g. stents with geometries in the hundreds of micrometres, this distinction is critical. All of the studies reported in the current literature compute fatigue crack growth parameters for Nitinol from the behaviour of cracks that are physically long (>1 mm), whereas many commercial devices have component sections that can be an order of magnitude smaller. If damage tolerant based design and life prediction strategies are to be used for small component medical devices, this use of large crack data must be acknowledged as a possible limitation that may lead to non-conservative predictions. Not only are small crack thresholds lower in most metallic materials, but at comparable stress intensities, small crack growth rates are significantly faster than those of corresponding large cracks, especially at the lower growth rates which dominate the life. Unfortunately, the measurement of small crack growth rates can be both difficult and highly prone to significant scatter. Therefore, we would not advocate the use of damage tolerant analyses as the primary life prediction methodology for small component medical devices at this time, particularly those manufactured from Nitinol, where the small crack experimental database is essentially non-existent.

Summary of trends in fatigue crack growth data

Much debate has surrounded the role of the transformation on the fatigue crack growth behaviour in Nitinol. To help resolve the differences in conclusions reported in the literature, we have combined the results of every crack growth experiment reported in the literature (see Table 3) and plotted the crack growth parameters m and ΔK_{th} versus a number of variables. Some important trends emerge from this comparison^{***}. The most important trend is the role of the initial microstructural

^{***} For the purposes of this comparison, the use of the term 'trend' simply indicates an observed tendency in the data. The use of 'significant' indicates a statistically significant ($p < 0.05$) trend as determined by a two tailed unpaired t test.



37 Crack growth parameters (Paris exponent m and fatigue threshold ΔK_{th}) for Nitinol as a function of R ratio. As expected, the threshold values decrease with increasing R ratio. Data are presented as mean \pm one standard deviation

phase in influencing the fatigue threshold stress intensity. Specifically, there is a significant increase in the fatigue threshold stress intensities in thermal martensite compared with all other phases, and a decreasing trend is noticeable towards the stable austenite condition. Given that these tests were conducted under stress controlled conditions and that threshold conditions correspond to nominal linear-elastic conditions, this observation is consistent with those data from total life studies discussed above and illustrated in Fig. 30b (Fig. 36). Conversely, the stable crack growth rate exponent m is virtually unaffected by the structural phase, R ratio, product form and specimen thickness. These trends are important for the design of fatigue resistant Nitinol devices. Indeed, they suggest that superelastic Nitinol may be more fatigue resistant when the operating temperature is between A_s and A_f (as opposed to above A_f but below M_d) since the material will still have superelastic properties, but will be closer in microstructure to the more fatigue crack growth resistant thermal martensite rather than stable austenite phase. Above A_s but below A_f , the existing martensite will be deformed upon application of stress. In addition, the existing volume fraction of austenite will transform into stress induced martensite. As such, the microstructure will become stabilised and inherit the characteristics of the thermal martensite, as discussed by MM1 for total life data.¹³

Not surprisingly, threshold stress intensities decrease significantly with increasing R ratio (Fig. 37). Although

Table 2 Small fatigue crack threshold stress intensity range estimations extrapolated from the corresponding large crack data for superelastic austenite Nitinol; after Robertson and Ritchie⁴³

R	$\Delta K_{th}/\text{MPa m}^{1/2}$	
	Large crack (experimentally determined)	Small crack (estimated from extrapolation line)
0.1	2.48	1.33
0.5	2.02	0.81
0.7	1.15	0.71

Table 3 Compilation of all fatigue crack growth data for Nitinol presented in the literature*

Phase	Reference	Temperature/°C				Product				Form		Anneal		Type		Dimension/ mm		Test conditions		Crack growth parameters			
		M_f	M_s	A_s	A_f	Test	ΔT	Composition/ at-%Ni	Plate	Bar	Tube	Strip	HR	CD	Time/ min	Temperature/ °C	B	X_f	v / Hz	R	m	C_f	ΔK_{th} / MPa m ^{1/2}
Mart	64	-55	-27	14	31	-190	-221	49.9	X					35	500	9	41	10	0.1	5.0	1.2×10^{-13}	5.0	
	64	-55	-27	14	31	-65	-96	49.9	X					35	500	9	41	10	0.1	3.7	5.0×10^{-12}	2.9	
	72	70	80	100	120	30	-90	49.9		X				60	750	13	51	20	0.1	3.4	
	72	70	80	100	120	30	-90	49.9	X		X			60	750	13	51	20	0.9	1.3	
	70	22	37	79	96	22	-73	...						60	750	10	...	50	0.1	3.2	7.8×10^{-12}	4.5	
	13	...	47	...	87	20	-67	49.9	X					60	950	10	60×90	2.3	<0.1	3.5	5.0×10^{-11}	...	
	72	70	80	100	120	60	-60	49.9	X		X			60	750	13	51	20	0.1	3.6	2.2×10^{-12}	3.8	
	72	70	80	100	120	60	-60	49.9	X		X			60	750	13	51	20	0.5	2.8	
	72	70	80	100	120	60	-60	49.9	X		X			60	750	13	51	20	0.7	2.5	
	72	70	80	100	120	60	-60	49.9	X		X			60	750	13	51	20	0.9	1.5	
	71	...	-7	6	38	-20	-58	50.8	X		X	X		60	400	1.5	32	8.5	1.5×10^{-17}	...	
	74	35	49	66	76	20	-56	49.4	X		X			60	930	10	50	15	0.1	4.5	9.4×10^{-10}	5.8	
Mix	79	63	20	-43	50.9	X		X	X		90	350	5	27	50	0.1	3.1	2.4×10^{-10}	0.83	
	72	70	80	100	120	80	-40	49.9	X		X	X		60	750	13	51	20	0.1	3.1	7.4×10^{-12}	3.6	
	72	70	80	100	120	80	-40	49.9	X		X	X		60	750	13	51	20	0.9	1.2	
	13	...	20	...	60	20	-40	49.7	X		X			60	950	10	60×90	2.3	<0.1	3.5	4.0×10^{-11}	2.4	
	79	50	20	-30	50.9		X				5	32	50	0.1	3.3	1.3×10^{-11}	2.5	
	70	-1	17	47	50	22	-28	10	...	50	0.1	3.5	1.6×10^{-10}	2.2	
	72	70	80	100	120	100	-20	49.9	X		X			60	750	13	51	20	0.1	3.2	9.0×10^{-12}	2.9	
	72	70	80	100	120	100	-20	49.9	X		X			60	750	13	51	20	0.9	1.5	
	79	39	20	-19	50.9	X		X	X		90	350	5	32	50	0.1	3.7	9.6×10^{-12}	2.6	
	71	...	31	19	44	26	-18	50.6	X		X	X		60	400	1.5	32	14	1.2×10^{-22}	...	
	70	-76	-26	15	37	22	-15	...	X		X			10	...	50	0.1	2.7	2.0×10^{-11}	1.7	
	71	...	-7	6	38	26	-12	50.8	X		X	X		60	400	1.5	32	4.8	1.5×10^{-12}	...	
	79	30	20	-10	50.9	X		X	X		90	300	5	32	50	0.1	2.8	3.3×10^{-11}	2.4	

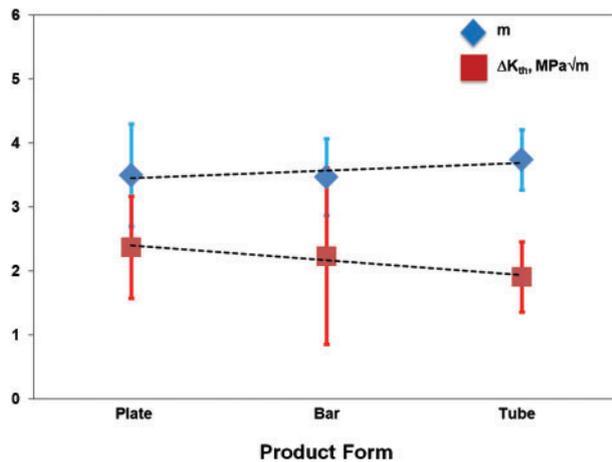
Table 3 Continued

Phase	Reference	Temperature/°C				Product										Crack growth parameters									
		M_f	M_s	A_s	A_f	Test	ΔT	Composition/ at-%Ni	Type				Form		Anneal Time/ min		Temperature/ °C	Dimension/ mm		Test conditions	Crack growth parameters				
		ΔK_{th}	M_f	M_s	A_s				A_f	Plate	Bar	Tube	Strip	HR				CD	B			X^\dagger	v	R	m
SE	72	70	80	100	120	120	0	49.9	X								60	750	13	51	20	0.1	3.4	1.2×10^{-11}	2.7
	72	70	80	100	120	120	0	49.9	X								60	750	13	51	20	0.5	2.0
	72	70	80	100	120	120	0	49.9	X								60	750	13	51	20	0.7	2.0
	72	70	80	100	120	120	0	49.9	X								60	750	13	51	20	0.9	1.9
	71	...	-12	-6	26	26	0	51.0	X								60	400	1.5	32	3.5	1.9×10^{-10}	...
	79	20	20	0	50.9	X								90	550	5	27	50	0.1	4.1	4.7×10^{-12}	2.6
	75	-16	4	12	33	37	4	50.6	X								0.2	5	15	0.2	3.0	1.5×10^{-10}	3.0
	79	15	20	5	50.9	X								90	550	5	32	50	0.1	3.5	1.0×10^{-11}	2.7
	63	-55	-27	14	31	37	6	49.9	X								35	500	9	41	10	0.1	3.0	1.2×10^{-10}	2.0
	64	-55	-27	14	31	37	6	49.9	X								35	500	9	41	10	0.1	3.0	7.6×10^{-11}	2.0
	64	-55	-27	14	31	37	6	49.9	X								35	500	9	41	10	0.5	3.1	1.6×10^{-10}	1.1
	64	-55	-27	14	31	37	6	49.9	X								35	500	9	41	10	0.7	3.0	3.6×10^{-10}	1
	71	...	-7	6	38	45	7	50.8	X								60	400	1.5	32	3.9	1.8×10^{-11}	...
	65	-100	-47	15	28	37	9	50.8	X								0.4	12	50	0.1	4.2	1.9×10^{-11}	2.5
	65	-100	-47	15	28	37	9	50.8	X								0.4	12	50	0.5	3.4	3.0×10^{-10}	1.9
	65	-100	-47	15	28	37	9	50.8	X								0.4	12	50	0.7	3.7	2.5×10^{-10}	1.4
	66	-100	-47	15	28	37	9	50.8	X								0.4	12	50	0.1	4.4	1.9×10^{-11}	2.5
	66	-100	-47	15	28	37	9	50.8	X								0.4	12	50	0.5	3.8	1.6×10^{-10}	2.0
	66	-100	-47	15	28	37	9	50.8	X								0.4	12	50	0.7	3.6	2.5×10^{-10}	1.2
	66	-100	-47	15	28	37	9	50.8	X								0.4	12	1	0.5	3.0	2.8×10^{-10}	...
	79	0	20	20	50.9	X								5	27	50	0.1	2.8	5.7×10^{-10}	0.58
	71	...	-7	6	38	60	22	50.8	X								60	400	1.5	32	3.7	7.8×10^{-11}	...
	79	-5	20	25	50.9	X								90	300	5	27	50	0.1	3.3	2.8×10^{-10}	0.65
	72	70	80	100	120	150	30	49.9	X								60	750	13	51	20	0.9	1.9
	72	70	80	100	120	150	30	49.9	X								60	750	13	51	20	0.1	4.2	5.5×10^{-12}	2.9
	70	-64	-52	-15	-13	22	35	...	X								10	...	50	0.1	2.9	2.4×10^{-12}	5.4
	71	...	-7	6	38	100	62	50.8	X								60	400	1.5	32	3.7	7.8×10^{-11}	...
Aust	64	-55	-27	14	31	120	89	49.9	X								35	500	9	41	10	0.1	3.4	4.0×10^{-11}	2.5
	64	-55	-27	14	31	120	89	49.9	X								35	500	9	41	10	0.5	3.5	7.7×10^{-11}	1.6
	64	-55	-27	14	31	120	89	49.9	X								35	500	9	41	10	0.7	3.8	8.4×10^{-11}	1.5
	71	...	-7	6	38	150	112	50.8	X								60	400	1.5	32	3.7	5.5×10^{-11}	...

*Data are sorted by $\Delta T = (A_f - T)$, where T is the test temperature. Areas corresponding to thermal martensite (Mart), mixed phase (Mix), superelastic austenite (SE) and stable austenite (Aust) are demarcated. For the purposes of statistical trend analysis (Figs. 36-39), the Miyazaki *et al.*⁷¹ data were excluded due to the questionable results as discussed earlier.

†X is the global size of the specimen, either diameter for disc specimens, or side length for square specimens.

‡Units = m/cycle/(MPa m^{1/2})².



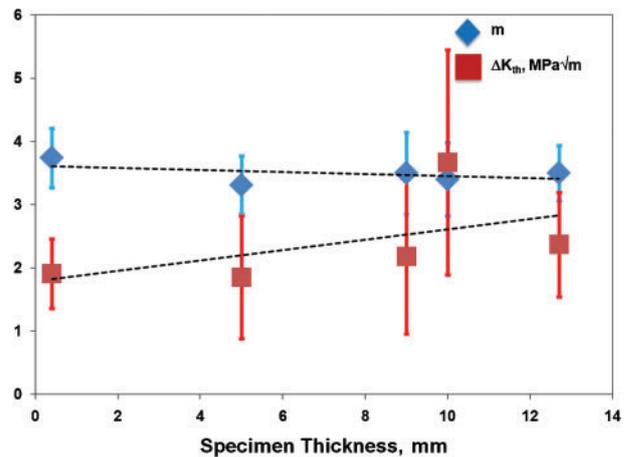
38 Crack growth parameters (Paris exponent m and fatigue threshold ΔK_{th}) for Nitinol in general were seemingly insensitive to the product form. Data are presented as mean \pm one standard deviation

Robertson and Ritchie⁶⁶ demonstrated the importance of evaluating the fatigue crack growth properties in the applicable product form to capture any crystallographic texture effects, the combined literature data in Fig. 38 show no significant effect among the various product forms (bar, plate and tube). Despite this lack of a trend in the overall data that show independent crack growth rates and thresholds, we would strongly advocate testing in appropriate product forms since those product specific tests may reveal vulnerabilities to fracture along geometry specific directions. Additionally, the role of plane stress versus plane strain conditions can be evaluated by plotting the crack growth parameters versus specimen thickness (Fig. 39). Although no experiments in the literature were conducted in Nitinol under truly plane strain conditions, we can reliably consider that the thinner specimens approach plane stress conditions whereas the thicker ones were mixed mode. The wide scatter in these data showed no specific trends in fatigue crack growth as a function of specimen thickness, similar to observations in other metallic materials.⁸¹ However, we would again advocate testing product specific geometries to accurately quantify the fatigue properties for each application of Nitinol material.

The principal conclusions that can be deduced from our survey of the fatigue crack propagation resistance of Nitinol alloys are that properties are: highly sensitive to structure (with martensite showing the best resistance and superelastic austenite the worst), highly sensitive to load ratio (increasing R leads to lower fatigue thresholds), insensitive to cyclic frequency (over the range 1–50 Hz in simulated physiological environments) and sensitive to texture and product form.

Fracture toughness

There have been very few experimental studies on the fracture toughness K_c of Nitinol. As listed in Table 4, there are only six papers in the literature addressing this important property.^{66,72,74,82–84} Most of the published data appear to be determined under non-plane strain conditions, specifically where the thickness of the test specimens was not large enough compared with the



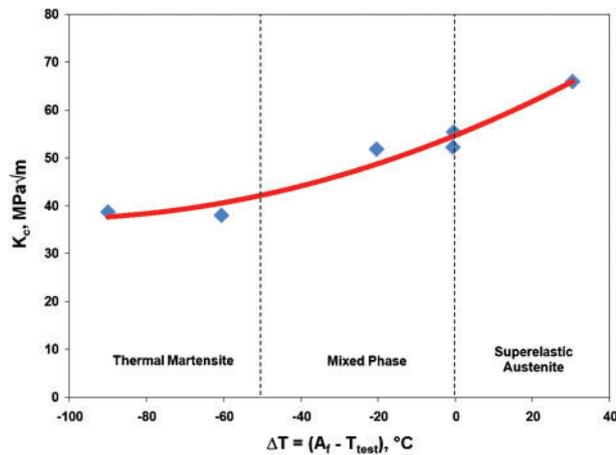
39 Crack growth parameters (Paris exponent m and fatigue threshold ΔK_{th}) for Nitinol as a function of specimen thickness. The fatigue threshold trend line shows an increase as the samples gain more through thickness constraint. However, if the outlying data from the 10 mm thick samples are excluded, the threshold values appear insensitive to sample thickness. Data are presented as mean \pm one standard deviation

plastic zone size, i.e. did not exceed $2.5(K_c/\sigma_y)^2$, where σ_y is the yield stress^{†††}; the one exception being a measurement on a higher strength stable austenitic alloy where the yield strength exceeds 650 MPa.⁸⁴ Furthermore, several of these reported measurements are questionable. For example, one study measured the toughness with a Nitinol tensile bar with a machined blunt notch rather than sharp precrack.⁸² Another group determined the fracture toughness from the unstable growth phase of a fatigue crack growth curve which again is a non-standard method for evaluating K_c values.⁷⁴ The fact that few ‘valid’ plane strain measurements have been made means that many of the reported toughness values for Nitinol are likely to be crack size dependent and geometry dependent. Despite this fact, they do represent quantifiable engineering K_c values that are pertinent to specific product forms.

The first literature citation of fracture toughness was reported by Holtz *et al.*⁷² in conjunction with their fatigue crack growth study. They conducted standard fracture toughness tests^{†††} in 12.7 mm thick hot rolled Nitinol plate material over a variety of temperatures such that the toughness as a function of stable microstructural phase was determined (Fig. 40). Below the M_f temperature, i.e. where martensite is the stable phase such that there is no superelasticity or phase

††† In thermal martensitic Nitinol, the detwinning strength (which indicates a divergence from linear-elastic behaviour) is typically <200 MPa. For superelastic Nitinol there is no true yield strength but rather a stress required for transformation from martensite to austenite, which is typically on the order of 500 MPa. The true (dislocation) yield strength for thermal martensitic and superelastic austenite occurs at much greater strains ($\sim 10\%$) and is usually in excess of 900 MPa.

††† A ‘standard’ fracture toughness tests herein refers to those that nominally comply with ASTM standards, specifically ASTM Standards E1820-08 (and formerly E399).⁸⁵ These tests typically involve a fatigue precracked compact tension or single edge notched bend specimen, with in-plane dimensions (crack size and width) roughly an order of magnitude larger than the crack tip plastic zone size to satisfy small scale yielding conditions, where fracture propagates from a monotonically loaded, atomically sharp flaw. If additionally plane strain conditions are required, the out of plane thickness of the specimen must also be order of magnitude larger than the plastic zone size.



40 Steady state fracture toughness of Nitinol as a function of stable phase showing an upward trend from thermal martensite to superelastic austenite; after Holtz *et al.*⁷²

transformation, the fracture toughness was lowest at $\sim 40 \text{ MPa m}^{1/2}$. An increase in fracture toughness to $53 \text{ MPa m}^{1/2}$ was measured when the specimens were raised into a mixed phase condition. A further increase in the toughness to $65 \text{ MPa m}^{1/2}$ was observed when the material was in the superelastic austenite form (150°C , which is 30°C above A_f) where the transformation to martensite could occur in response to the stresses ahead of the crack tip.

Holtz *et al.*⁷² concluded that the increasing toughness values were caused by a different mechanism of fracture at higher temperatures, citing the observation of a flatter fracture surface in the superelastic austenite sample. However, examination of their micrographs revealed fatigue striations on the superelastic Nitinol sample. Accordingly, we believe that the authors drew their conclusions erroneously from the fatigue precracked region, rather than the overload fracture region, of the fracture surface. It is our opinion that, in actuality, the fracture mechanisms were essentially identical in each test – martensite either preexists or is stress induced immediately ahead of the growing crack tip. Consequently, we would proffer an alternative explanation for the increase in fracture toughness with increasing temperature. We first note that the observed toughness increase occurred only after the testing temperature was raised above A_s (i.e. into the mixed phase region). Second, as the transformation (i.e. plateau) stress increases with increasing differential between the A_s and test temperature, a higher applied load would be required to initiate martensite ahead of the crack tip. Since the toughness is directly proportional to this load^{§§§}, it follows that fracture toughness will increase as the required transformation loads rises with increasing temperature above A_s . Furthermore, the toughness values measured by Holtz *et al.* are all significantly higher than those measured in other martensitic and superelastic Nitinol fracture studies. We are of the opinion that since the samples used in this study were fully annealed (750°C , 1 h), the grain size

§§§ The fracture toughness is determined from $K_c = Q\sigma_{app}(\pi a)^{1/2}$, where σ_{app} is the applied stress, Q is a geometric factor on the order of unity and a is the crack length (or half length).

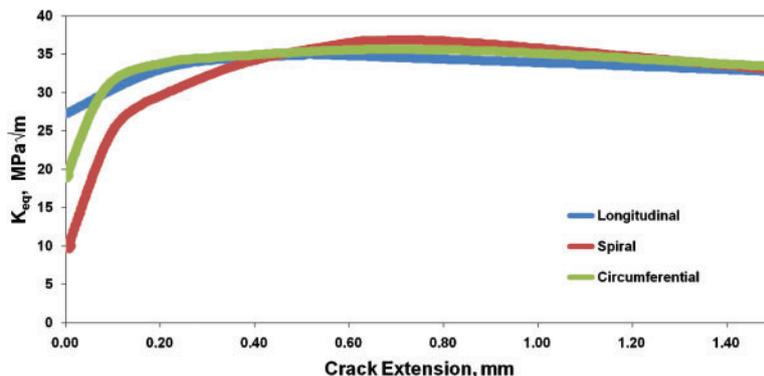
(most likely $\sim 100 \mu\text{m}$) was relatively large compared with contemporary as thermomechanically processed material ($\sim 10 \mu\text{m}$). At these large grain sizes, there would be very few grains in the transformation zone ahead of the growing crack tip, which could certainly distort the toughness values measured.

He *et al.*⁸² studied the role of hydrogen embrittlement on the fatigue and fracture of Nitinol. It is unknown what product form of Nitinol that these researchers used, but it was clear from the text that the material was superelastic austenite at the test temperature (20°C). They computed a fracture toughness value of $39.2 \pm 2.8 \text{ MPa m}^{1/2}$ in the virgin (non-embrittled) form. They further found that the toughness logarithmically decayed with increased hydrogen content as per: $K_c = 78.8 - 9.8 \ln c_t$, where c_t is the total hydrogen concentration (atomic hydrogen + hydrides) measured between 500 and 4200 wppm. It should be noted that these researchers used a highly non-standard specimen geometry for determining fracture toughness, specifically, a machined notched, 20 mm gauge length tensile dogbone specimen with gauge dimensions $0.5 \times 3.0 \text{ mm}$. Correspondingly, although the authors refer to their measured values as (mode I) fracture toughnesses, a more appropriate term would be ‘notch toughness’ as their samples did not contain a nominally atomically sharp crack. Their toughness values are almost certainly artificially high because the blunt notch would significantly lessen the intensity of the stress field there. Typically, the measured (apparent) fracture toughness scales with roughly the square root of the notch root radius.⁸⁶

Using three very different product forms, respectively bar, plate and tube ranging in thickness from 0.4 to 10 mm, Vaidyanathan *et al.*,⁷⁴ Gollerthan *et al.*^{83,84} and Robertson and Ritchie⁶⁶ each measured nearly identical fracture toughness values of $\sim 30 \text{ MPa m}^{1/2}$ for superelastic austenite. Although the fatigue crack growth behaviour was affected significantly by the presence of TiC inclusions, the fracture toughness values calculated from the unstable growth portion of fatigue crack growth curves by Vaidyanathan *et al.*⁷⁴ were similar regardless of inclusion density (0–20% inclusion volume content)^{¶¶¶}. Robertson and Ritchie⁶⁶ presented the only fracture toughness values for Nitinol in the form of rising R curve^{****} behaviour, which indicates the stress intensity as a function of initial crack extension necessary to initiate and sustain subcritical cracking (Fig. 41). Furthermore, they demonstrated a marked influence of the crystallographic texture on the initiation toughness, but no crystallographic influence over steady state growth. Specifically, toughness values at the onset of crack extension ranged from 10 to $27 \text{ MPa m}^{1/2}$, with

¶¶¶ Although it is recognised that the unstable growth regime of a fatigue crack growth curve is dominated by the K_{max} value in the fatigue cycle approaching the fracture toughness, determination of the fracture toughness from such data cannot necessarily account for the role of cyclic deformation, crack tip plasticity, crack size effects and loading rate. Therefore, Vaidyanathan *et al.*'s fracture toughness data should be viewed with some caution.

**** The crack resistance or R curve provides an assessment of the fracture toughness in the presence of subcritical crack growth. It involves measurements of the crack driving force, e.g. K , G or J , as a function of crack extension Δa . The value of the driving force at $\Delta a \rightarrow 0$ provides a measure of the crack initiation toughness, whereas the slope or the maximum value of the R curve can be used to characterise the crack growth toughness.



41 Rising R curve behaviour observed in thin walled (0.4 mm thick) superelastic austenite Nitinol tube specimens. K_{eqr} represents the equivalent stress intensity to cause cracking. Also observed was dependence in fracture toughness on the crack growth direction which highlights the influence of texture in this product form; after Robertson and Ritchie⁶⁶

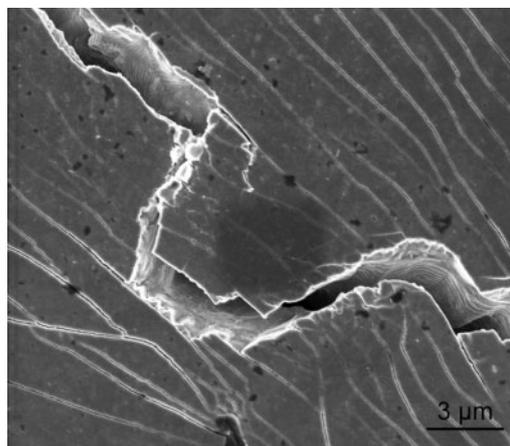
the lowest fracture toughness occurring 45° ('spiral') to the drawing direction and highest parallel ('longitudinal') to the drawing direction.

Finally, Gollerthan *et al.*⁸⁴ examined the fracture of Nitinol using *in situ* synchrotron X-ray diffraction experiments, in particular to investigate the nature of the structure of the transformation zone surrounding the crack tip. Earlier researchers⁶⁴ suggested that the formation of stress induced martensite could be suppressed ahead of a growing crack in a thick (plane strain) specimen due to constraint from the high triaxial (hydrostatic) stress field to explain trends in their fatigue crack growth data. However, Gollerthan *et al.*'s diffraction experiments give evidence to refute that hypothesis. Specifically, their data provided definitive crystallographic evidence of fully reversible stress induced martensite ahead of a growing crack, which has also been observed experimentally under non-plane strain conditions in thin walled flattened tube.⁷⁸ These observations provide some understanding why widely varying Nitinol product forms and microstructures (i.e. thermal martensite versus superelastic austenite) all display virtually identical steady state fracture toughness values of $\sim 30 \text{ MPa m}^{1/2}$. Indeed, since crack growth behaviour is most influenced by the microstructure in the immediate vicinity of the crack tip, it is perhaps not surprising that the fracture behaviour of a thermal martensite specimen is nominally the same as a superelastic austenite specimen with stress induced martensite at the crack front. The mechanism of fracture in Nitinol changes only when the temperatures are sufficiently high (above M_d) such that the microstructure into which the crack grows is stable austenite rather than martensite. Gollerthan *et al.*'s⁸⁴ results show that the fracture toughness values increase dramatically ($53 \text{ MPa m}^{1/2}$) for stable (non-transforming) austenite, which we would suggest is associated with the increase in yield strength concomitant with an increase in temperature above A_s .

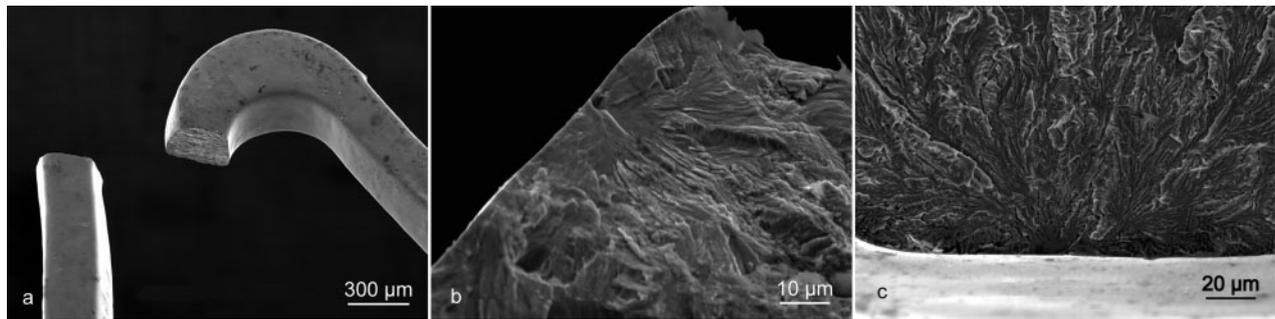
Fractography

Despite its unique mechanical behaviour, the morphology of Nitinol's fracture and fatigue surfaces is quite similar to conventional engineering metals. The laboratory created fatigue specimens shown in Fig. 42 show many important fracture characteristics for Nitinol in the superelastic austenite condition. First, the crack

propagation path is noticeably tortuous corresponding to local variances in grain texture that either promote or suppress the phase transformation, and consequently the yield strength and fatigue resistance of individual grains. Next, notice that stable crack growth is possible, as indicated by the presence of microcracks alongside the primary fracture path that serve to dissipate energy and enhance fatigue life. Lastly, notice the clearly defined fatigue striations on the primary crack surface which are observed in many low cycle or medium cycle fatigue surfaces; these indicate the local cycle by cycle progress of the crack during fatigue crack propagation. Striations in high cycle fatigue specimens are often absent because they are severely worn (or even completely obliterated) due to the long life of stable crack growth and/or spaced so fine that only high magnification microscopes are capable of resolving them.



42 Superelastic austenite Nitinol ($A_f=25^\circ\text{C}$) laboratory created stress controlled ($R=0.1$) fatigue fracture in 37°C HBSS. The crack morphology is characterised by path tortuosity corresponding to grain texture, stable microcrack formation and striations on the primary fracture surface (see inside the crack). Striation spacings are such that the crack was propagating at $\sim 10^7 \text{ m/cycle}$, implying that at this location, crack growth was occurring in the upper portion of the Paris law regime near the K_{max} dominated unstable growth regime. Crack growth is from top left to bottom right in the photograph. After Robertson and Ritchie⁶⁶



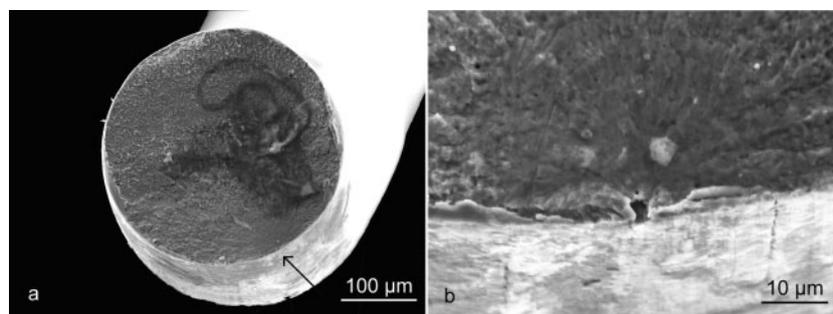
43 Stent-like diamond Nitinol fatigue specimen with laboratory created fatigue fracture showing *a* overview of fracture location, *b* fatigue origination from a below critical flaw sized inclusion and *c* fatigue origin in base metal (non-inclusion) region under the same fatigue loading conditions. Note the fatigue striations typical of uniaxial strain controlled fatigue conditions; after Pelton *et al.*⁴⁶

As might be expected, fatigue fractures in Nitinol originate in regions of local stress/strain concentration. This often coincides with surface, or near surface, inclusions. Although fatigue crack origination at inclusions is observable both in bench top experiments (Fig. 43) and commercial devices (Fig. 44), neither the presence nor abundance of inclusions has been shown to have a major deleterious effect on the fatigue life (e.g. *see* Fig. 10). To the contrary, as discussed earlier, when the volume fraction and size of inclusions conform to ASTM standards for wrought Nitinol,⁴² they are below the critical flaw size for outright fracture; accordingly, fatigue lives of Nitinol alloys with normal versus extra low inclusion content are essentially the same.^{35,36,38}

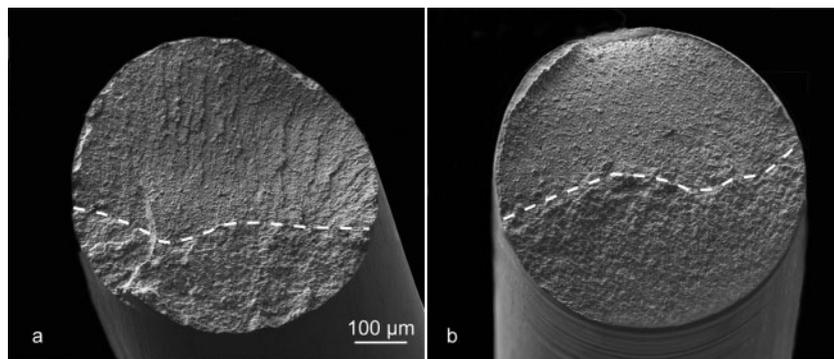
Since fatigue of both thermal martensite and superelastic austenite material requires crack propagation through a martensitic phase, one would expect similarities between the fracture surfaces. Indeed, Figs. 45 and 46 show (laboratory created) rotary bending fatigue fracture surfaces in these two phase conditions tested under 1 and 10% alternating strain conditions. As expected, both phases produced a rough fracture topography in the low cycle (high alternating strain) condition (Figs. 45*b* and 46*b*). Indeed, 10% strain conditions lead to unstable crack growth with fatigue crack propagation through only a small percentage of the cross-section before complete ductile overload failure (see line delimiting the regions). Conversely, under the high cycle 1% alternating strain condition, the

thermal martensite sample (Fig. 45*a*) is able to sustain $\sim 50\%$ stable crack growth over $\sim 10^7$ cycles resulting in a relatively flat fracture through the sample before ductile overload fracture ensued. The same 1% alternating strain input to the superelastic austenite material (Fig. 46*a*) resulted in medium cycle ($\sim 10^5$) fatigue fracture. Interestingly, the superelastic austenite samples tested at both 1 and 10% alternating strains show similar surface roughness typical of the fast fatigue fracture and/or final ductile overload event. Because the two samples have significantly different fatigue lifetimes but similar fracture surfaces, it suggests that the majority of the fatigue lifetime in these superelastic austenitic Nitinol samples is spent in crack nucleation rather than stable crack growth.

Single event overload fractures in Nitinol are distinctly different from their fatigue fractured counterparts. Overload fracture surfaces are characterised by necking, ductile dimpling and microvoid coalescence. This high energy fracture usually leads to rougher fracture morphologies than fatigue fractured specimens, with the exception of the near terminus area of a fatigue fractured component which is subject to the same ductile overloading conditions as monotonically loaded specimens. Figures 47 and 48 respectively show that the overload fracture surfaces in thermal martensite and superelastic austenite are quite similar. Figure 49 presents an overload fractured commercial superelastic austenite intravascular device that failed while in service. Notice the distinctive necking and ductile dimples



44 *a* a low magnification of a commercial intravascular superelastic austenite wire based Nitinol implant that fractured *in vivo* and was subsequently explanted from the patient. This relatively flat fracture surface is indicative of high cycle bending fatigue with an origin at a near surface inclusion. *b* The 2 µm inclusion is well below the critical flaw size ($15\text{--}50\text{ }\mu\text{m}$)⁴³ for superelastic austenitic Nitinol and did not contribute to the failure, but did serve as a local stress/strain discontinuity from which the crack nucleated. Notice the worn fatigue clamshell pattern near the fracture origin and ductile overload at the terminus that are classical of fatigue fractures. The darkened discoloured region in the upper right quadrant of the fracture is adherent biological material on the explanted specimen and is not part of the native material. Images after Hull and Robertson⁸⁹



- 45 *a* laboratory created low cycle rotary bending fatigue fracture of thermal martensitic Nitinol showing primarily unstable crack growth ($\sim 25\%$ penetration) before complete ductile overload fracture. The rough surface is comparable with high energy (low cycle) fatigue crack surfaces in conventional metals. Conditions: 0.6 mm diameter black oxide wire, $R=-1$, 10% strain amplitude, room temperature, $A_f=70^\circ\text{C}$, 1600 cycles to failure. *b* laboratory created high cycle rotary bending fatigue fracture of thermal martensitic Nitinol showing stable crack growth through $\sim 50\%$ of the cross-section. The relatively flat fracture surface is comparable with low energy (high cycle) fatigue crack surfaces in conventional metals. Conditions: 0.6 mm diameter black oxide wire, $R=-1$, 1% strain amplitude, room temperature, $A_f=70^\circ\text{C}$, $\sim 10^7$ cycles to failure. Images after Wick *et al.*³⁷ with the demarcation line separating the fatigue fracture (below line) from the ductile overload fracture (above line)

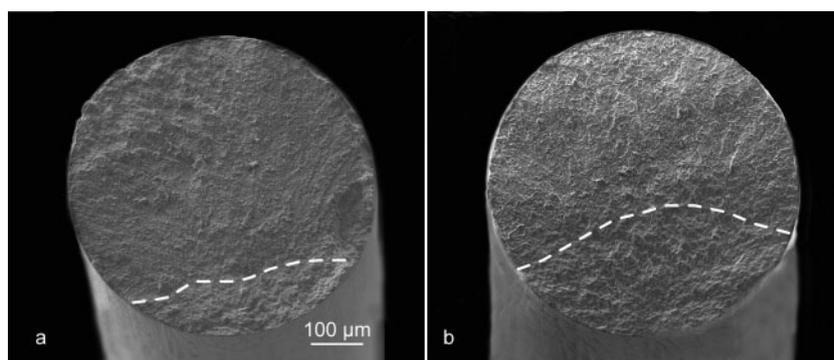
that are indicative of a Nitinol overload fracture as compared with the flatter, low energy fatigue fracture of a different commercial device shown in Fig. 44.

Discussion and concluding remarks

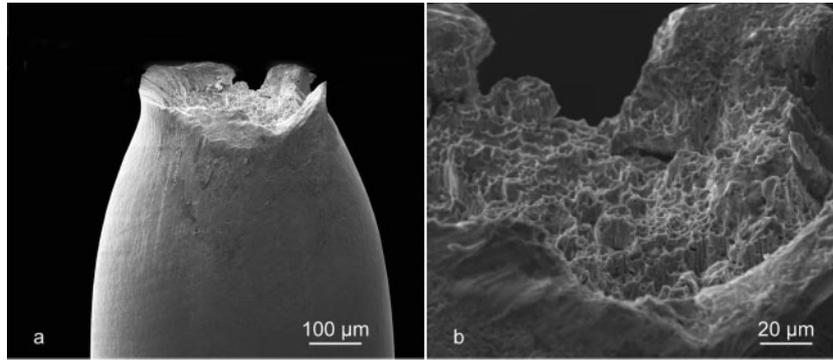
It is without question that the issue of cyclic fatigue resistance is of extreme importance to the design and safe application of medical implant devices manufactured from Nitinol alloys. Indeed, medical devices manufactured from any material that are expected to survive millions of cyclic deformations over their lifetime require scrutiny of the fatigue and fracture resistance. The phenomenon of fatigue is the root cause of many premature fractures in any metallic component, and biomedical implants are certainly no exception. Biomedical devices manufactured from Nitinol (e.g. peripheral vascular stents) can experience up to 40 million loading cycles each year. Therefore, these implants need to be designed

to survive some 10^8 – 10^9 fatigue cycles without failure over the lifetime of the patient. However, as noted above, by definition, such biomedical devices tend to be geometrically small with many critical components, e.g. the struts of a stent, sized well under a millimetre. This limitation in physical size can complicate the choice of a methodology that is most approach for their fatigue design and for the prediction of their safe life.

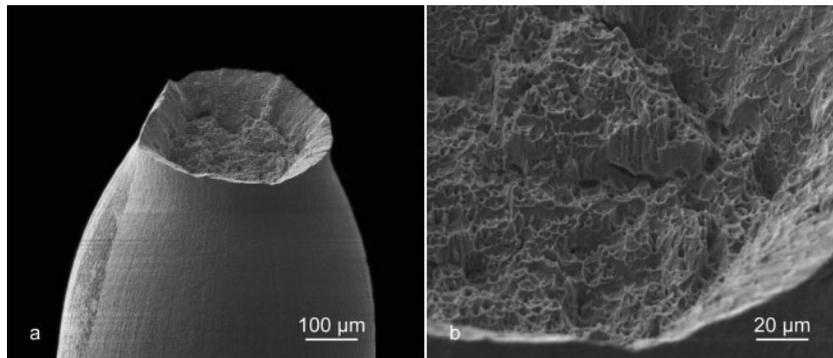
Traditionally, fatigue design and life prediction procedures have been conducted using total life analyses, either involving stress–life or strain–life data. The utilisation of stress–life versus strain–life data requires an understanding of the cyclic boundary conditions. Consider the cyclic loading due to the human gait on various places in the anatomy. A stent placed in a popliteal artery may experience cyclic bending deformations of 90° in response to knee flexion and therefore require strain based methods. Conversely, a hip replacement will be subjected to cyclic loads proportional to the



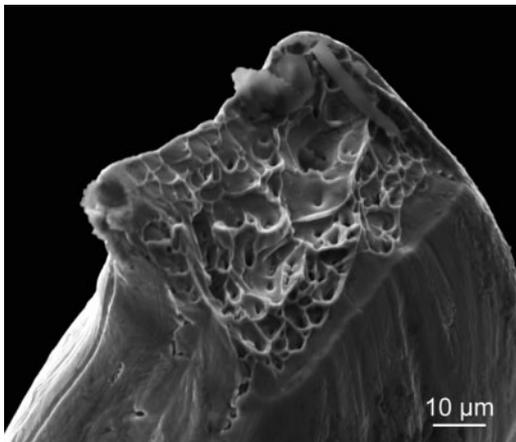
- 46 *a* laboratory created low cycle fatigue fracture surface of superelastic austenitic Nitinol with very little ($<5\%$) unstable crack growth followed by high energy ductile overload (rough morphology). Conditions: 0.6 mm diameter oxide free wire, $R=-1$, 10% strain amplitude, room temperature, $A_f=10^\circ\text{C}$, 554 cycles to failure. *b* laboratory created medium cycle fatigue fracture surface of superelastic austenitic Nitinol with limited ($\sim 10\%$) stable crack growth into the wire before ductile overload fracture. Notice that the rough morphology is comparable with the low cycle fatigue specimen *a* suggesting that the majority of fatigue life is spent in nucleation rather than propagation. Conditions: 0.6 mm diameter oxide free wire, $R=-1$, 1% strain amplitude, room temperature, $A_f=10^\circ\text{C}$, 11,448 cycles to failure. Images after Wick *et al.*³⁷ with the demarcation line separating the fatigue fracture (below line) from the ductile overload fracture (above line)



47 *a* laboratory created monotonic tensile overload specimen of thermal martensitic Nitinol showing classic necking with subsequent 'cup and cone' fracture common in ductile metals. *b* higher magnification fractograph of the thermal martensitic Nitinol sample shown in Fig. 45a. Conditions: 0.6 mm diameter black oxide wire, 2.5 mm min^{-1} strain rate, room temperature, $A_f=70^\circ\text{C}$



48 *a* laboratory created monotonic tensile overload sample of superelastic austenitic Nitinol showing a similar necked and 'cup and cone' fracture surface to the thermal martensite sample shown in Fig. 45. *b* higher magnification fractograph of the superelastic austenite Nitinol sample shown in Fig. 49a. Conditions: 0.6 mm diameter oxide free wire, 2.5 mm min^{-1} strain rate, room temperature, $A_f=10^\circ\text{C}$



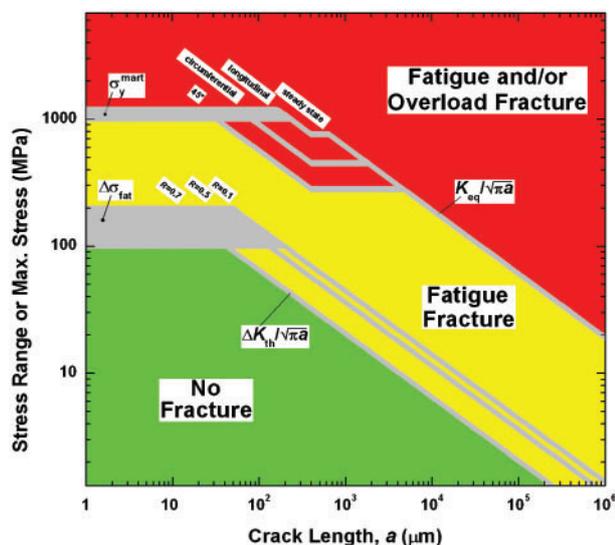
49 Commercial superelastic austenitic Nitinol intravascular device that fractured in service due to single event, high energy ductile overload failure. Notice the necking of this laser cut tubular product and dimpling on the fracture surface that is indicative of classical ductile overload fracture

weight of the patient in response to each footstep and hence require a stress based approach. In the current commercial landscape, nearly all biomedical implants that utilise Nitinol are in deformation/strain controlled environments. As Nitinol's usage broadens into the orthopaedic environment or non-medical applications,

this trend may shift to include more stress controlled conditions.

Constant life diagrams based upon total life experiments, such as those presented in Figures 20 and 22 are essential in understanding the material characteristics specific to Nitinol. These plots are created using test (coupon) specimens cycled under a variety of deformation conditions. This approach relies on the fatigue of large numbers of specimens in order to produce a statistically viable sampling to provide quantitative material properties. As it is generally impractical to repeat these tests for every iterative design, instead, we would encourage the use of the appropriate literature data as a guideline for evaluating material design of Nitinol medical devices. Device specific testing is subsequently performed on a much smaller number of samples after the design has passed the 'litmus' test of the literature constant life data for Nitinol (e.g. pulsatile fatigue of stents in Fig. 23). Specifically, 'fatigue to life' experiments using commercially manufactured components such that the thermomechanical history and surface characteristics (e.g. presence of flaws) are representative of the end product are generally required to demonstrate device specific resistance to fatigue loading conditions that accurately simulate the duty cycle and life of a product.

Damage tolerant approaches relying on fracture mechanics based crack growth data have been proposed



50 Robertson *et al.* combined data from multiples researchers into this diagram that unifies total life, damage tolerant and fracture toughness of Nitinol in the superelastic austenite condition into a single design tool. Crack length corresponds to the maximum expected flaw in the material, which can be equated to the minimum surface flaw detectable during visual inspection procedures. The maximum stress intensity specifically addresses single event excursions in the deformation or load input to the device such as those experienced during crimping of a stent into a catheter. Lastly, the stress intensity range coincides with the alternating cyclic fatigue boundary conditions experienced by the Nitinol device during its lifetime. Reprinted with permission from Robertson and Ritchie⁴³

and successfully applied to larger implant devices such as metallic and pyrolytic carbon mechanical heart valve prostheses.^{87,88} However, this approach has not yet found utility in the current commercial landscape of Nitinol cardiovascular implants. As the damage tolerant methodology is based solely on the crack propagation life, it assumes that no time is spent in crack initiation and is thus inherently conservative in predicting safe lifetimes, which can then be presented as a function of the size of any such preexisting flaws. This approach will become important in Nitinol applications where the dimensions of the devices are on the order of centimetres such as orthopaedic implants such that the implant can withstand stable crack growth without complete failure. For now, however, the usefulness of the damage tolerant approach for physically small (smallest feature ≤ 1 mm) Nitinol devices is in quantifying a critical flaw size, below which fatigue crack growth can be considered dormant. This critical flaw size may be used for quality control procedures to establish inspection pass/failure criterion for use during visual inspection of parts for surface defects.

To appreciate the utility of these different approaches, Robertson and Ritchie constructed a useful diagram (Fig. 50) that unified the total life, damage tolerant and fracture toughness data into a single design tool for determination of product susceptibility to overload and/or fatigue fracture.⁴³ This diagram plots the maximum expected flaw size (which corresponds to visual

inspection acceptance criteria) versus the maximum stress intensity (corresponding to single event overloads such as crimping of a stent into its delivery catheter) and alternating stress intensity (corresponding to the cyclic loads experienced while in service). The diagram, however, is still somewhat limited in its usefulness since it is presented in terms of stress rather than strain. Indeed, superelasticity of Nitinol manifests as non-unique stress conditions when the specimen experiences stresses/strains anywhere on the upper or lower plateau. However, its benefit is as a design tool for evaluating superelastic austenite material specific designs since it combines the data from multiple researchers into a single diagram.

Acknowledgements

Support for SWR and ARP was provided by Nitinol Devices and Components (NDC), Fremont, CA. The involvement of ROR was funded by the Director, Office of Science, Office of Basic Energy Sciences, Division of Materials Sciences and Engineering, of the US Department of Energy under contract no. DE-AC02-05CH11231.

References

1. K. Otsuka: 'Introduction to the R-phase transition', in 'Engineering aspects of shape memory alloys', (ed. T. W. Duerig *et al.*), 36–45; 1990, London, Butterworth-Heinemann.
2. K. Otsuka and X. Ren: 'Physical metallurgy of Ti–Ni-based shape memory alloys', *Prog. Mater. Sci.*, 2005, **50**, 511–678.
3. T. W. Duerig, K. N. Melton, D. Stöckel and C. M. Wayman: 'Engineering aspects of shape memory alloys'; 1990, London, Butterworth-Heinemann.
4. K. Bhattacharya: 'Microstructure of martensite', 288; 2003, Oxford, Oxford University Press.
5. T. Duerig, A. Pelton and D. Stoeckel: 'An overview of nitinol medical applications', *Mater. Sci. Eng. A*, 1999, **A273–A275**, 149–160.
6. D. Stöckel, A. R. Pelton and T. Duerig: 'Self-expanding Nitinol stents: material and design considerations', *Eur. Radiol.*, 2004, **14**, 292–301.
7. A. R. Pelton, J. DiCello and S. Miyazaki: 'Optimisation of processing and properties of medical-grade Nitinol wire', *Min. Invas. Ther. Allied Technol.*, 2000, **9**, 107–118.
8. A. Wöhler: 'Über die Festigkeitsversuche mit Eisen und Stahl', *Zeitschrift Bauwesen*, 1870, **20**, 73–106.
9. L. F. Coffin, Jr and J. F. Tavernelli: 'The cyclic straining and fatigue of metals', *Trans. Am. Inst. Min. Eng.*, 1959, **215**, 794–807.
10. S. S. Manson: 'Thermal stress and low cycle fatigue', 125–192; 1966, New York, McGraw-Hill.
11. J. D. Morrow: 'Cyclic plastic strain energy and fatigue of metals', in 'Internal friction, damping, and cyclic plasticity', 45–87; 1965, Philadelphia, PA, ASTM.
12. P. C. Paris and F. Eredogan: 'A critical analysis of crack propagation laws', *J. Bas. Eng. (Trans. ASME Ser. D)*, 1961, **85**, 528–534.
13. K. N. Melton and O. Mercier: 'Fatigue of NiTi thermoelastic martensites', *Acta Metall.*, 1979, **27**, (1), 137–144.
14. K. N. Melton and O. Mercier: 'The effect of the martensitic phase transformation on the low cycle fatigue behaviour of polycrystalline Ni–Ti and Cu–Zn–Al alloys', *Mater. Sci. Eng.*, 1979, **40**, (1), 81–87.
15. S. Miyazaki, Y. Kohiyama, K. Otsuka and T. W. Duerig: 'Effects of several factors on the ductility of the Ti–Ni alloy', *Mater. Sci. Forum*, 1990, **56–58**, 765–770.
16. S. Miyazaki, T. Imai, Y. Igo and K. Otsuka: 'Effect of cyclic deformation on the pseudoelasticity characteristics of Ti–Ni alloys', *Metall. Trans. A*, 1986, **17A**, 115–120.
17. S. Miyazaki: 'Thermal and stress cycling effects and fatigue properties of NiTi alloys', in 'Engineering aspects of shape memory alloys', (ed. T. W. Duerig *et al.*), 394–413; 1990, London, Butterworth-Heinemann.

18. L. C. Brinson, I. Schmidt and R. Lammering: 'Stress-induced transformation behavior of a polycrystalline NiTi shape memory alloy: micro and macromechanical investigations via in situ optical microscopy', *J. Mech. Phys. Solids*, 2004, **52**, (7), 1549–1571.
19. D. M. Norfleet, P. M. Sarosi, S. Manchiraju, M. F. X. Wagner, M. D. Uchic, P. M. Anderson and M. J. Mills: 'Transformation-induced plasticity during pseudoelastic deformation in Ni-Ti microcrystals', *Acta Mater.*, 2009, **57**, (12), 3549–3561.
20. N. Ono and A. Sato: 'Plastic deformation governed by the stress induced martensitic transformation in polycrystals', *Trans. Jpn Inst. Met.*, 1988, **29**, 267–273.
21. N. Ono, A. Satoh and H. Ohta: 'A discussion on the mechanical properties of shape memory alloys based on a polycrystal model', *Trans. Jpn Inst. Met.*, 1989, **30**, (10), 756–764.
22. S. Miyazaki, Y. Sugaya and K. Otsuka: 'Effects of various factors on fatigue life of Ti-Ni alloys', Proc. MRS Int. Meet. on 'Advanced materials', Tokyo, Japan, May–June 1988, MRS, Vol. 9, 251–256.
23. A. R. Pelton, T. Duerig and D. Stöckel: 'A guide to shape memory and superelasticity in nitinol medical devices', *Min. Invas. Ther. Allied Technol.*, 2004, **13**, (4), 218–221.
24. A. R. Pelton, J. DiCello and S. Miyazaki: 'Optimization of processing and properties of medical-grade nitinol wire', Proc. Int. Conf. on 'SMST', Pacific Grove, CA, USA, April–May 2000, International Organization on SMST, 361–374.
25. K. N. Melton and O. Mercier: 'Fatigue of NiTi and CuZnAl shape memory alloys', Proc. 5th Int. Conf. on 'The strength of metals and alloys', (ed. P. Haasen), 1243–1248; 1979, Oxford, Pergamon Press.
26. J. J. Lewandowski and P. Lowhaphandu: 'Effects of hydrostatic pressure on mechanical behaviour and deformation processing of materials', *Int. Mater. Rev.*, 1998, **43**, (4), 145–187.
27. T. Buchheit and J. Wert: 'Modeling the effects of stress state and crystal orientation on the stress-induced transformation of NiTi single crystals', *Metall. Mater. Trans. A*, 1994, **25A**, (11), 2383–2389.
28. T. Buchheit and J. Wert: 'Predicting the orientation-dependent stress-induced transformation and detwinning response of shape memory alloy single crystals', *Metall. Mater. Trans. A*, 1996, **27A**, (2), 269–279.
29. K. Gall and H. Sehitoglu: 'The role of texture in tension-compression asymmetry in polycrystalline NiTi', *Int. J. Plast.*, 1999, **15**, (1), 69–92.
30. K. Gall, H. Sehitoglu, Y. I. Chumlyakov and I. V. Kireeva: 'Tension-compression asymmetry of the stress-strain response in aged single crystal and polycrystalline NiTi', *Acta Mater.*, 1999, **47**, (4), 1203–1217.
31. M. Barney, D. Xu, S. W. Robertson, V. Schroeder, R. O. Ritchie, A. R. Pelton and A. Mehta: 'Impact of thermomechanical texture on the superelastic response of nitinol implants', *J. Mech. Behav. Biomed. Mater.*, 2011, **4**, 1431–1439.
32. S. S. Manson and G. R. Halford: 'Fatigue and durability of structural materials'; 2006, Materials Park, OH, ASM International.
33. Y. Kim and S. Miyazaki: 'Fatigue life of Ti-50 at.% Ni and Ti-40Ni-10Cu (at.%) shape memory alloy wires', Proc. 2nd Int. Conf. on 'SMST', Asilomar, CA, USA, March 1997, International Organization on SMST, 473–478.
34. S. Miyazaki, K. Mizukoshi, T. Ueki, T. Sakuma and Y. Liu: 'Fatigue life of Ti-50 at.% Ni and Ti-40Ni-10 Cu (at.%) shape memory alloy wires', *Mater. Sci. Eng. A*, 1999, **A273–A275**, 658–663.
35. M. Reinhoehl, D. Bradley, R. Bouthot and J. Proft: 'The influence of melt practice on final fatigue properties of superelastic NiTi wires', Proc. Int. Conf. on 'SMST', Pacific Grove, CA, USA, April–May 2000, International Organization on SMST, 397–403.
36. J. Sheriff, A. R. Pelton and L. A. Pruitt: 'Hydrogen effects on nitinol fatigue', Proc. ASM Materials & Processes for Medical Devices Conf., Minneapolis, MN, USA, August 2004, ASM International, 38–43.
37. A. Wick, X.-Y. Gong, J. Fino, J. Sheriff and A. R. Pelton: 'Bending fatigue characteristics of nitinol', Proc. ASM Materials & Processes for Medical Devices Conf., Minneapolis, MN, USA, August 2004, ASM International, 15–20.
38. N. Morgan, A. Wick, J. DiCello and R. Graham: 'Carbon and oxygen levels in nitinol alloys and the implications for medical device manufacture and durability', Proc. Int. Conf. on 'SMST', Pacific Grove, CA, USA, May 2006, ASM International, 821–828.
39. B. T. Berg: 'Bending of superelastic wires: Part I – experimental aspects', *J. Appl. Mech.*, 1995, **62**, 459–466.
40. A. Wick, O. Vöhringer and A. Pelton: 'The bending behavior of NiTi', *J. Phys. IV*, 1995, **5**, 789–794.
41. S. Suresh: 'Fatigue of materials'; 1998, Cambridge, Cambridge University Press.
42. 'Standard specification for wrought nickel-titanium shape memory alloys for medical devices and surgical implants', F2063-05, ASTM International, West Conshohocken, PA, USA, 2005.
43. S. W. Robertson and R. O. Ritchie: 'A fracture-mechanics-based approach to fracture control in biomedical devices manufactured from superelastic Nitinol tube', *J. Biomed. Mater. Res. Part B: Appl. Biomater.*, 2008, **84**, 26–33.
44. T. W. Duerig, D. E. Tolomeo and M. Wholey: 'An overview of superelastic stent design', *Min. Invas. Ther. Allied Technol.*, 2000, **9**, (3/4), 235–246.
45. T. W. Duerig: 'The use of superelasticity in modern medicine', *MRS Bull.*, Feb. 2002, 101–104.
46. A. R. Pelton, V. Schroeder, M. R. Mitchell, X.-Y. Gong, M. Barney and S. W. Robertson: 'Fatigue and durability of nitinol stents', *J. Mech. Behav. Biomed. Mater.*, 2008, **1**, 153–164.
47. A. R. Pelton: 'Nitinol fatigue: a review of microstructures and mechanisms', *J. Mater. Eng. Perform.*, 2010, **20**, (4–5), 613–617.
48. G. R. Zadno and T. W. Duerig: 'Linear and non-linear superelasticity in NiTi', Proc. MRS Int. Meeting on 'Advanced materials', Tokyo, Japan, May–June 1988, MRS, 201–206.
49. R. M. Tabanli, N. K. Simha and B. T. Berg: 'Mean stress effects on fatigue of NiTi', *Mater. Sci. Eng. A*, 1999, **A273–A275**, 644–648.
50. R. M. Tabanli, N. K. Simha and B. T. Berg: 'Mean strain effects on the fatigue properties of superelastic NiTi', *Metall. Mater. Trans. A*, 2001, **32A**, (7), 1866–1869.
51. D. Tolomeo, S. Davidson and M. Santinoranout: 'Cyclic properties of superelastic nitinol: design implications', Proc. Int. Conf. on 'SMST', Pacific Grove, CA, USA, April–May 2000, International Organization on SMST, 409–417.
52. N. B. Morgan, J. Painter and A. Moffat: 'Mean strain effects and microstructural observations during *in vitro* fatigue testing of NiTi', Proc. Int. Conf. on 'SMST', Pacific Grove, CA, USA, May 2003, International Organization on SMST, 303–310.
53. A. R. Pelton, X.-Y. Gong and T. W. Duerig: 'Fatigue testing of diamond-shaped specimens', Proc. Int. Conf. on 'SMST', Pacific Grove, CA, USA, May 2003, International Organization on SMST, 293–302.
54. M. R. Mitchell: 'Fundamentals of modern fatigue analysis for design', in 'ASM handbook', 'Fatigue and fracture', 227–262; 1996, Materials Park, OH, ASM International.
55. E. Macherauch: 'Praktikum in werkstoffkunde'; 2002, Wiesbaden, Friedrich Vieweg & Sohn Verlagsgesellschaft mbH.
56. N. Rebelo and M. Perry: 'Finite element analysis for the design of nitinol medical devices', *Min. Invas. Ther. Allied Technol.*, 2000, **9**, (2), 453–462.
57. N. Rebelo, M. Hsu and H. Foadian: 'Simulation of superelastic alloys behavior with ABAQUS', Proc. Int. Conf. on 'SMST', Pacific Grove, CA, USA, April–May 2000, International Organization on SMST, 457–469.
58. A. Runciman, D. Xu, A. R. Pelton and R. O. Ritchie: 'An equivalent strain/Coffin-Manson approach to multiaxial fatigue and life prediction in superelastic Nitinol medical devices', *Biomaterials*, 2011, **32**, (22), 4987–4993.
59. E. Patoor, P. Barbe, A. Eberhardt and M. Berveiller: 'Internal stress effect in the shape memory behaviour', *J. Phys. IV*, 1991, **5**, 95–100.
60. Y. Liu and P. G. McCormick: 'Thermodynamic analysis of the martensitic transformation in NiTi. I. Effect of heat treatment on transformation behaviour', *Acta Metall. Mater.*, 1994, **42**, (7), 2401–2406.
61. P. G. McCormick and Y. Liu: 'Thermodynamic analysis of the martensitic transformation in NiTi. II. Effect of transformation cycling', *Acta Metall. Mater.*, 1994, **42**, (7), 2407–2413.
62. K. N. Smith, P. Watson and T. H. Topper: 'A stress-strain function for the fatigue of metals', *J. Mater. JMLSA*, 1970, **5**, (4), 767–778.
63. A. L. McKelvey and R. O. Ritchie: 'Fatigue-crack propagation in Nitinol, a shape-memory and superelastic endovascular stent material', *J. Biomed. Mater. Res. Part A*, 1999, **47A**, (3), 301–308.
64. A. L. McKelvey and R. O. Ritchie: 'Fatigue-crack growth behavior in the superelastic and shape-memory alloy Nitinol', *Metall. Mater. Trans. A*, 2001, **32**, (3), 731–743.
65. J. M. Stankiewicz, S. W. Robertson and R. O. Ritchie: 'Fatigue-crack growth properties of thin-walled superelastic austenitic Nitinol tube for endovascular stents', *J. Biomed. Mater. Res. Part A*, 2007, **81A**, (3), 685–691.

66. S. W. Robertson and R. O. Ritchie: 'In vitro fatigue-crack growth and fracture toughness behavior of thin-walled superelastic Nitinol tube for endovascular stents: a basis for defining the effect of crack-like defects', *Biomaterials*, 2007, **28**, 700–709.
67. J. L. Proft, K. N. Melton and T. W. Duerig: 'Transformation cycling of Ni–Ti and Ni–Ti–Cu shape memory alloys', Proc. MRS Int. Meet. on 'Advanced materials', Tokyo, Japan, May–June 1988, MRS, 159–164.
68. A. Mehta, X.-Y. Gong, V. Imbeni, A. R. Pelton and R. O. Ritchie: 'Understanding the deformation and fracture of nitinol endovascular stents using in situ synchrotron X-ray micro-diffraction', *Adv. Mater.*, May 2007, **19**, 1183–1186.
69. R. V. Marrey, R. Burgermeister, R. B. Grishaber and R. O. Ritchie: 'Fatigue and life prediction for cobalt-chromium stents: a fracture mechanics analysis', *Biomaterials*, 2006, **27**, (9), 1988–2000.
70. R. H. Dauskardt, T. W. Duerig and R. O. Ritchie: 'Effects of *in situ* phase transformation on fatigue-crack propagation in titanium–nickel shape-memory alloys', Proc. MRS Int. Meet. on 'Advanced materials', Tokyo, Japan, May–June 1988, MRS, Vol. 9, 243–249.
71. S. Miyazaki, M. Suizu, K. Otsuka and T. Takashima: 'Effect of various factors on fatigue crack propagation rate in Ti–Ni alloys', Proc. MRS Int. Meet. on 'Advanced materials', Tokyo, Japan, May–June 1988, Vol. 9, 263–268.
72. R. L. Holtz, K. Sadananda and M. A. Imam: 'Fatigue thresholds of Ni–Ti alloy near the shape memory transition temperature', *Int. J. Fatigue*, 1999, **21**, S137–S145.
73. A. Ohta, N. Suzuki and T. Mawari: 'Effect of Young's modulus on basic crack propagation properties near the fatigue threshold', *Int. J. Fatigue*, 1992, **14**, (4), 224–226.
74. R. Vaidyanathan, D. C. Dunand and U. Ramamurty: 'Fatigue crack-growth in shape-memory NiTi and NiTi–TiC composites', *Mater. Sci. Eng. A*, 2000, **A289**, (1–2), 208–216.
75. P. Filip, M. Paliwal and K. Mazanec: 'Fatigue crack propagation in pseudoelastic TiNi smart microcomponents', *Zeitschrift Metall*, 2004, **95**, 356–361.
76. S. Gao and S. Yi: 'Experimental study on the anisotropic behavior of textured NiTi pseudoelastic shape memory alloys', *Mater. Sci. Eng. A*, 2003, **A362**, (1–2), 107–111.
77. S. W. Robertson, X. Y. Gong and R. O. Ritchie: 'Effect of product form and heat treatment on the crystallographic texture of austenitic Nitinol', *J. Mater. Sci.*, 2006, **41**, (3), 621–630.
78. S. W. Robertson, A. Mehta, A. R. Pelton and R. O. Ritchie: 'Evolution of crack-tip transformation zones in superelastic Nitinol subjected to in situ fatigue: A fracture mechanics and synchrotron X-ray microdiffraction analysis', *Acta Mater.*, 2007, **55**, (18), 6198–6207.
79. K. Gall, J. Tyber, G. Wilkesanders, S. W. Robertson, R. O. Ritchie and H. J. Maier: 'Effect of microstructure on the fatigue of hot-rolled and cold-drawn NiTi shape memory alloys', *Mater. Sci. Eng. A*, 2008, **A486**, (1–2), 389–403.
80. R. O. Ritchie and J. Lankford: 'Small fatigue cracks: a statement of the problem and potential solutions', *Mater. Sci. Eng.*, 1986, **84**, 11–17.
81. R. O. Ritchie, R. F. Smith and J. F. Knott: 'Effects of thickness on fibrous fracture from a notch and on fatigue-crack propagation in low-strength steel', *Met. Sci.* 1975, **9**, (11), 485–492.
82. J. Y. He, K. W. Gao, Y. J. Su, L. J. Qiao and W. Y. Chu: 'The role of hydride, martensite and atomic hydrogen in hydrogen-induced delayed fracture of TiNi alloy', *Mater. Sci. Eng. A*, 2004, **A364**, (1–2), 333–338.
83. S. Gollerthan, D. Herberg, A. Baruj and G. Eggeler: 'Compact tension testing of martensitic/pseudoplastic NiTi shape memory alloys', *Mater. Sci. Eng. A*, 2008, **A481–A482**, 156–159.
84. S. Gollerthan, M. L. Young, A. Baruj, J. Frenzel, W. W. Schmahl and G. Eggeler: 'Fracture mechanics and microstructure in NiTi shape memory alloys', *Acta Mater.*, 2009, **57**, (4), 1015–1025.
85. 'Standard test method for measurement of fracture toughness', E1820-08, ASTM International, West Conshohocken, PA, USA, 2008.
86. R. O. Ritchie, B. Francis and W. L. Server: 'Evaluation of toughness in AISI 4340 alloy steel austenitized at low and high temperatures', *Metall. Trans. A*, 1976, **7A**, (6), 831–837.
87. R. O. Ritchie and P. Lubock: 'Fatigue life estimation procedures for the endurance of a cardiac valve prosthesis: stress/life and damage-tolerant analyses', *J. Biomech. Eng. Trans. ASME*, 1986, **108**, 153–160.
88. R. O. Ritchie: 'Fatigue and fracture of pyrolytic carbon: a damage-tolerant approach to structural integrity and life prediction in 'ceramic' heart-valve prostheses', *J. Heart Valve Dis.*, 1996, **5**, (Suppl. 1), S9–S22.
89. J. E. Hull and S. W. Robertson: 'Bard recovery filter: evaluation and management of vena cava limb perforation, fracture, and migration', *J. Vasc. Interv. Radiol.*, 2009, **20**, (1), 52–60.