Unsettling standards: the biological age controversy

Abstract: One key debate within the sociology of aging and the life-course in the past decade has been focused on understanding the extent to which life course processes have become de-standardised. However, little attention has been devoted to the infrastructural processes that would support such a transition. This paper focuses on one of such standards, ‘biological age’, and the 40 year controversy around the method and purpose of its measurement. Drawing on published research, archival and interview data collected in Europe and North America, the paper suggests that the persistent uncertainty surrounding biological age measurements is structured by differing interlocking relationships between normative ideals of the life-course and methodological approaches to knowledge making to understand and manage the relationship between aging, health and illness.
In May 2011, popular science and mainstream papers reported on a rift between the two winners of the 2009 Nobel Prize for Medicine, Elizabeth Blackburn and her former assistant Carol Greider. While the prize had been won for work they had done together in the 1980s to understand the function of telomeres - structures on the end of chromosomes, which regulate cell replication -, their disagreement hinged on the use of telomere length as a measure of biological or ‘real’ age of individuals (Pollack 2011). Blackburn, a professor of biology at the University of California, San Francisco, had recently launched a company, Telome Health, providing telomere length testing as a gauge of individuals’ “underlying health” and risk of developing “diseases of aging.” (Blackburn in Dickinson 2011) Greider, now a professor of molecular biology at Johns Hopkins University, questioned the evidence-base of her former colleague’s claims, arguing that “the science really isn’t there to tell us what the consequences are of your telomere length.” (Greider in Pollack 2011) While Blackburn was confident enough about the validity and reliability of the technique to plan to sell it as a “biomarker of aging”, Greider was less assured of the scientific meaning of telomere length, arguing that using just the test “one wouldn’t be able to tell how old the person really is.” (Greider in Pollack 2011)

Their disagreement fueled public concerns about the commercialization of genetic testing, the possible uses of this type of information by insurance companies, and the legal frameworks that regulate the release and access to such tests. These were further compounded when, in 2012, a class-action lawsuit was filed accusing Telomerase Activation Sciences (TA Sciences) – another company selling telomere length testing – of deception by promoting a proprietary herbal extract – TA65 – as capable of reversing the effects of aging (Borrel 2012). Counteracting these concerns, proponents of the tests, likened telomere length measuring to currently available, routine techniques such as cholesterol assessment, and emphasized that tests were being provided through qualified physicians “who understand and interpret the results for their patients.” (Matlin in Connor 2012) They also argued that most of the interventions thought to affect telomere length were within the domain of health maintenance rather than through the application of medical technologies. Despite these assurances, critics still argued that the “use of telomere length to aid diagnosis or risk assessment in unselected individuals is premature.” (von Zglinicki 2012:344) as salient, basic questions about the normal variance of telomere length in human
populations and its relation to biological aging processes remain unanswered (Sanders and Newman 2013).

Focused on a technical aspect of contemporary molecular biology, this controversy reached wider public attention because of the possible implications of reliable telomere testing for the management of the aging process for individuals and populations. Often conceived as a combination of genetic and environmental/lifestyle factors through which individuals become, at different rates, increasingly susceptible to illness, aging has, in the past three decades, been institutionally aligned with policies and programs that emphasize the role of preventive biomedicine and health maintenance practices (e.g. Butler et al 2008) as key to the maximization of older people’s participation in the polity and the economy (Moreira and Palladino 2008). This, in turn has sustained a renewed scientific interest in developing and validating measures of individual biological aging rate, or biomarkers of aging.

A biomarker of aging is defined as “a biological parameter of an organism that either alone or in some multivariate composite will, in the absence of disease, better predict functional capability at some late age, than will chronological age.” (Baker and Sprott 1988:28) As the definition suggests, measures of biological aging aim to replace chronological age as an index of the functional ability or health of individuals, promising a ‘personalized’ assessment. Biomarkers of aging would be useful, advocates argue, not only to assess an individual’s susceptibility to developing illness but also, and importantly, to evaluate the effects of emerging age-retarding interventions (Butler et al 2004). Thus, since the 1980s a variety of funding initiatives, research programs and consortia have propelled the number of publications on biomarkers of aging from just below 10 between 1985 and 1991 to around 250 in 2011 alone (Web of Knowledge data July 2012). However, despite such investment and interest, the search for a method to measure biological age (BA) has been characterized by limited success, and, as the telomere controversy illustrates, is concomitant to a sustained and on-going uncertainty about the purpose, validity, accuracy, reliability, and practicality of proposed aging rate measures.

The purpose of this paper is to explore and understand the persistent uncertainty surrounding the purpose, validation and implementation of BA measurements. It asks, what are the institutional, normative, methodological and material factors that have shaped the controversy over BA in the past 4 decades? It addresses this question by
drawing on the approach to technical controversies developed in the sociology of science and technology to analyze documentary, archival and interview data relating to the controversy on BA. Analysis of this data suggests that the disagreement around biomarkers of aging is structured by differences in the way in which understandings of measurement and standardization are linked to institutional configurations of expertise in the management of aging, health and illness. In particular, I suggest that while BA advocates propose to differentiate between individuals by deploying ranking standards that distally embody the technical expertise of biometricians, sceptics prefer to emphasize that instrumented measurements of aging and health must always rely on the close pastoral guidance of biomedical experts. I argue that these two stances on aging standards have interacted to reinforce normative commitments for and against the social and political organization of contemporary biomedicine (Moreira and Palladino 2009).

In so doing, the paper addresses a crucial but seldom explored question in the sociology of aging and the life course. While sociologists of the life course have been, at least since the late 1980s, focused on understanding the extent to which there has been a shift from a reliance on chronological age to segment the life-course towards a de-standardized, individualized life-course in advanced economies, little attention has been devoted to the infrastructural processes that would support such a transition: what are the technologies, standards and conventions that would, in practice, equip a personalized management of the life-course? Within the social sciences, this neglect is not particular to the sociology of aging (Timmermans and Epstein 2010) but undermines our capacity to understand how age measurements are assembled and how they might become embedded in and transform social institutions.

A key consequence of casting the sociological eye on standards is that uncertainties, conflicts and contingencies that normally remain in the backstage of analysis come into full view (Lampland and Star 2009). But uncertainty is not intended here to have a negative meaning. Indeed, one could argue that uncertainty is an inherent quality of knowledge production or research (Latour 1998), which becomes particularly acute when the concepts, instruments and data analysis techniques themselves are disagreed upon by specialists or contested in wider society. In cases of public controversies, such as the one analyzed in this paper, uncertainties focus not only on the knowledge base for institutional rules for distributions of rights and responsibilities, but also on
the institutional character of knowledge making procedures themselves. It becomes a question not only of ensuring the production and dissemination of scientifically robust research, but also of understanding how knowledge production is organized, for what purposes, and for whose benefit. Drawing on such conceptual background, by focusing on the BA controversy, the paper aims to contribute to the understanding of the complex interweaving relationship between expertise, scientific and technological standards and social, normatively embedded age identities in contemporary societies.

After a brief methodological note, the paper outlines how growth in scientific interest in ‘personalized’ age measurements during the 20th century combined with changes in the institutional apparatus of standard making practices to produce a consensus around the purposes and methods of BA measurement. It then examines how, within this context, researchers critically engaged with proposed BA measures identifying key weaknesses in their development, validation and deployment, and how these critiques were, in turn, countered by advocates. The paper analyses how the process of controversy helped consolidate two divergent positions about the role of standards in the management of aging, health and illness, which have structured this domain of research, practice and policy until the present day.

Methodological note

The data analyzed in the paper was collected as part of an on-going research project on age standardization and is a combination of interview, archival and documentary material. This includes a database containing approximately 4000 references published between 1935 – the year marking the establishment of modern gerontology (Park 2008) - and 2013. A historical review of the literature selected by number of citations and relevance led to the identification of a series of chronologically overlapping case studies of research programs, institutions and approaches concerned with the concept and measurement of age for which archival material was collected. Semi-structured interviews with key actors (N=8) in the controversy, identified through the historical literature review, were conducted with the aim to enrich the understanding of the configurations of institutional, normative, conceptual and material resources underpinning different positions.

Drawing on the framework proposed in Moreira (2010), for this paper, the data were analyzed through the use of procedures of technical controversy mapping.
Controversy mapping has its origins in the sociology of scientific knowledge, where it was used to understand the factors that facilitated or hindered epistemic consensus within scientific communities, later developing into a technique to analyze the dynamic interaction between knowledge generation or technological innovation and social change (Latour 1987; Law and Bijker 1992; Latour 2005; Callon, Barthe and Lascoumes 2009; Venturini 2010). Drawing primarily on publicly available documents, it aims to identify cluster relationships between issues and organizational or institutional actors in relation to specific issues or domains of concern, and to describe and map the changes in such relationships in time. Discursive positions taken by actors in the public space are recursively analyzed through how their relative position about the issue at stake is affected by the temporal unfolding of the controversy itself. That is to say that rather than regarding debate as a form of regress to underlying assumptions and positions, technical controversy mapping looks for key turning points as transformative and generative.

Following this methodology in data analysis, entailed the use of a) historical methods to map main events – the turning points - of the controversy, b) standard qualitative analysis techniques of constant comparison, thematic coding and deviant case analysis to characterize the stances, and alliances formed by groups within particular stages of the controversy; and c) analysis of the institutional networks in which such repertoires and positions are embedded at different phases of the controversy. This work was aided by the use of NVivo software.

**Personalizing age measurement**

In the sociology of aging and the life course, it is usually agreed that widespread and consistent use of techniques of age measurement were key to establish the structural and cultural apparatus that segmented the ‘modern’ life course in three distinct stages (Mayer 2009). In this process, chronological age (CA) became the main criterion – the *marker* – for a modern stratified system of public rights and duties that included military draft, school enrollment and access to welfare rights such as pensions (Kohli 1986). This relied on changes to the *infrastructural apparatus* of classification and quantification of age by central agencies.

Of key significance was the convergence between the epistemic norms of statistical reasoning and the information requisites of state administration in the latter part of the
19th century (Desrosieres 1991), whereby a wide range of classification practices and decision making procedures inherent to modern bureaucracies became contingent on numerical rather than categorical age-reporting. However, the widespread use of CA was dependent also upon complex and continuous work of educating, policing and recording knowledge of age among the population (Treas 2009:74-76; also Beaud and Prevost 1994). These processes combined together to raise awareness of CA and its normative implications in European and North American societies in between the 19th and 20th centuries (Chudacoff 1989).

As I discussed above, over the past decade or so, a debate has emerged about the extent to which the forces of globalization, labor market de-regulation, re-structuration of public services and individualization processes have produced heterogeneity of transitions to adulthood and retirement, or increasingly de-institutionalized and/or de-standardized life course trajectories (Beck 2001; Bruckner and Meyer 2005; Hughes and Waite 2007; Kohli 2007; Philipson 2015). Curiously, less attention has been devoted to understanding the role of age measurement in this process (but see: Settersten and Mayer 1997; Katz and Marshall 2004; Bytheway 2011). Historical analysis of the scientific fields engaged in age measurement and standardization suggests that, during 20th century, epistemic and normative uncertainties about the validity of CA as a universal metric has combined with wider shifts in standardization processes in late modernity (Busch 2011), whereby measures and scales promise individualization and/or ‘personalization’ of technologies or services (Moreira submitted). In the remainder of this section, I explore how these two became interlinked.

The origins of the interrogation of CA as a measurement can be traced back to the consolidation of the sciences of growth in the US in the beginning of the 20th century. In that period, Progressive reforms of the American nation, and growing anxieties about modernization, induced the expansion of privately and publicly funded research on child development (Smuts 2008; Prescott 2004). These aimed to replace the previous concern with poor and delinquent children with a scientific understanding of the ‘normal’ child. Between the mid-1920s and the early 1930s, researchers in the emerging field of child development established a variety of studies aimed at examining ‘growth’ by means of serial observations of selected children, collecting
both biometrical data and psychosocial assessments (e.g.: the Harvard Growth Study, the Oakland Growth Study).

The concepts and ideals of the child development research movement can be seen as embodied in the figure of Lawrence Frank, one of the key planners of the movement. Trained as a social scientist, Frank was typically troubled by modernization and what he saw as the resulting, growing disjunction between habitual human behaviors and industrial, technological culture. He thus regarded the understanding of processes of normal physiological and psychological development as key to the design of beneficent social institutions and the management of individual behavior (Bryson 1998:410). As the manager of the Laura Spelman Rockefeller Memorial Fund (LSRMF), and as editorial board member of *Growth*, one of the key journals of the field, Frank construed his primary responsibility to be the promotion and dissemination of studies of the ‘whole child’, and in linking contemporaneous understandings of social problem to the scientific program of child development research. Significantly, as the Great Depression began to ravage the American society and the federal government became entangled in debates concerning the future Social Security Act (1935), Frank was again called upon to create an equivalent enthusiasm for research on the elderly as he had for children. As vice-president of the Josiah Macy Foundation, he sought to shape a nascent gerontological program through the idea that it was necessary to understand ‘normal, physiological’ aging to be able to design an effective response to the problems posed by the elderly (see Achenbaum 1995:76–79).

This resonated with the vision proposed by Vincent Cowdry, the main scientific instigator of the role of the Macy Foundation in sponsoring research into the ‘problems of the elderly’. As Park (2008) has documented, Cowdry, a cytologist working at Washington University, was concerned with the divergence between differentiating physiological processes at the level of the cell and tissue, and the increasing typification of older people as ‘useless’ with basis of their chronological age. With Frank’s assistance, in 1935, Cowdry gathered together a group of experts for a conference on aging, which can be seen as marking the establishment of modern gerontology (Park 2013), and which he later published as *Problems of Aging* (Cowdry 1939). In this forum, drawing on the ideas of Nobel Prize winning physiologist and eugenist Alexis Carrel, Cowdry proposed that rate of aging in tissues was determined
by their surrounding environment of nutrients, regardless of the organism’s chronological age. In this, Cowdry compounded the case against CA from a biological and medical perspective. In particular, he questioned the idea that there was an alignment between the calendar and the various ‘speeds’ at which different organs develop and decay. Cowdry’s proposal was to become central to modern gerontology as it aligned a program of social reforms with contemporaneous medical and scientific ideas and practices regarding growth and aging.

Frank’s role was also crucial in helping establish another key figure in modern gerontology, Nathan Shock. Shock had begun his academic career in the mid-1930s as a researcher in the Oakland Growth Study (see above), a study in which Frank had invested great hope towards the end of his tenure at the LSRMF (Moreira and Palladino 2011). Focusing on physiological changes in adolescence, Shock was able to establish that the onset of a physiological event - menarche - was more important than CA in structuring changes in development. It was thus no coincidence that, when advising the Federal government on the establishment of an intramural gerontology program at the National Institutes of Health, Frank recommended Shock as new chief of the Gerontology Unit, when the first chief, Edward Stieglitz, resigned for a personal reason in 1941 (Park 2013). There, Shock was able to establish, during the 1940s, a program of research that drew on his experience in measuring ‘physiological age’, now using institutionalized elderly subjects from the Baltimore Department of Public Welfare.

While it is clear that a critique of CA and a parallel interest in measures of physiological age matured in the years between the wars, it was only after 1945 that the idea of ‘biological age’ became explicitly articulated in the field of gerontology. In one of the first recorded uses of the concept, Harry Benjamin, an endocrinologist practicing in New York, nowadays mostly known for his work with Alfred Kinsey on transsexuality (Benjamin 1966), suggested that the aim of the concept of BA was to develop “for the individual what our actuaries and biostatisticians have figured out so ingeniously for groups.” (Benjamin 1947:226) This, he argued, was important for two purposes.

On the one hand, it would be possible to replace CA with ‘objective’ measurement of age when assessing individuals’ capacity to work and decision to retire. On the other hand, knowledge of individual BA could assist in monitoring and assessing the effects
of healthy living, medical interventions or “gerontotherapy, which indicates the
treatment of the aging process as such.” (Benjamin 1947:225) This last reason was of
particular significance to Benjamin, who had been involved providing hormonal
(‘gonadal’) treatments in his private practice since the 1920s. The importance of
Benjamin’s proposal relies in the fact that it encapsulates two key ideas within BA
research since the 1950s.

One relates to the purpose of the measurement, through the suggestion that it was
possible to provide an individualized assessment of age to maximize older people’s
health and well-being, through the use of health technologies, and the tailoring of
social or commercial services. Indeed, more than 30 years later, when explaining the
reasons behind organizing the first conference on Biological Markers of Aging, the
Director of the recently created National Institute of Aging, Robert Butler, argued
that,

[t]here are both scientific and socioeconomic imperatives for developing
biological makers of aging. The scientific imperatives derive from […] the
possibility that certain age-related phenomena […] may be controlled through
intervention, [the testing of which] is dependent upon accurate measures of
biological aging. The socioeconomic imperative [stem] from economic
perturbations that have threatened the integrity of the Social Security System, [and
which have motivated] proposals to increase the age of social security
eligibility. […] Because of the increased age of the workforce and conflicts over
retirement age […] , we must be able to assess properly the impact of aging on
human performance. (Butler 1982:vi)

In the intervening years, a variety of research projects and programs had been
established to ascertain BA by drawing on this combination of social and technical
‘imperatives’. Of key importance here had been the establishment of longitudinal
studies of aging such as the Duke Longitudinal Studies of Normal Aging (1955), the
Baltimore Longitudinal Study of Aging (BLSA) (1958), or the Atomic Bomb
Casualty Commission Adult Health Study (1957), the latter being a study which,
while seemingly focused exclusively on the effects of radiation on human aging, still
justified its wider relevance in terms of its “importance in industry.” (Hollingsworth,
Hashizume and Hablon 1965:12)
But perhaps the most significant case of the normative platform on which BA research lay was the BLSA. Specifically established as part of Nathan Shock’s attempt, as Director of the Gerontology Unit, to translate the individualist and voluntarist thrust of Eisenhower’s initiatives on aging (Lockett 1983:71–72) into a research program, the BLSA aimed to provide evidence for the wider institutional recognition of “older persons as individuals, not as a class, and their wide differences in needs, desires and capacities.” (NWS:21 ‘Progress in the field of gerontology’, 13 July 1956) This aim required methodological attention to the monitoring of change in individuals’ physiological characteristics and psychological capacities over time, rather than regressing these to the mean of the age group. This is linked to the second key idea in Benjamin’s 1947 paper, which relates to method.

Benjamin’s suggestion was that the assessment of BA was best done through assembling a battery of different metrics or ‘factors’. Similarly, in another seminal paper in this field, Irwin Mackay Murray (1951), an anatomy academic at New York State University, proposed that “a method that would in some way combine various physiologic functions would give a more useful assessment of the individual” than single measurements (Murray 1951:120). To do this, Murray, drawing on advice from the eminent statistician Ronald A. Fisher (Marks 2003), proposed the use of the multiple regression method to assess the differences between CA and BA of selected age-groups. Drawing on the consensus that different physiological systems would age at different rates (see above), the significance of Murray’s method is related, for our purposes, to how it implies an apparatus of data collection instruments and techniques on which statistical tools can be applied. In other words, from early in the development of BA measurement, it became clear that these ranking standards could only be developed with recourse to a complex arrangement of physiological data and statistical expertise.

More instructive from this perspective is Alex Comfort’s work on this subject. A key figure in the consolidation of biological gerontology, Comfort was mostly known for promoting an approach to the measurement of aging drawing on the synthesis of Darwinism and Mendelian genetics (Medawar 1952), and hinging on the effects of the ‘force of mortality’ upon survival curves (Comfort 1956:22-44). However, during the 1960s, in his capacity as editor of Experimental Gerontology, he became increasingly persuaded of the possibility of manipulating rates of senescence in cells
and organisms, in particular by replications of the work of Mckay on caloric restriction, by Harman’s research on anti-oxidants (AC 24: Correspondence with Denham Harman 1965-73) as well as new developments in pharmacological approaches to aging (Bender, Kormendy and Powell 1970). This, and the need to find an application for biology of aging within the problem-driven research policy environment of the 1960s (Ruivo 1994), led him to the view that “new attempts to work out battery tests of human physiological age [were] overdue [because they were] justified by experimental necessity.” (Comfort 1972:101)

Much of the focus of Comfort’s work on a ‘battery test’ related to how the correlation coefficients of different proposed batteries to the ‘curve of mortality’ balanced against the pragmatics of its implementation. Advocating “a procedure used for clinical screening […] based on a flow-type center using lay staff” to collect data (Comfort 1972:105) and computer facilities to store and calculate data, Comfort shows an insightful awareness of the infrastructural - technical, financial and sociological - requirements of BA measurement. In a methodological proposition that anticipates much of the criteria for validating biomarkers of aging (Kirkwood 1998), Comfort articulates an evaluative, regulatory role for the measurement of individual, biological age, the implementation of which relies on streamlined data collection and management, and the expertise of biometricians.

Comfort’s work also underlines the importance of the relationship between method and purpose. In this respect, Murray already had recognized, that the size of his sample - a small cross sectional population study- limited the generalizability of his conclusions but that “[s]tandards based on [a random] sample could be used in assessing applicants for life insurance [and] to patients in clinical medicine.” (Murray 1951:125) Equally, the relationship between the methodological design for, and the practical purposes of, BA measurement became a key aspect of discussion in the establishment of the longitudinal studies of aging referred to above. What specific metrics were going to be recorded and in what way were these to be combined or used in statistical analysis of results? Where the measurements drawn from the sample generalizable and applicable to the population as a whole? What policy, medical or public health needs would this information address?

As Paolo Palladino and I (2011) have shown, these questions were fundamental to the dynamic that sustained the BLSA in its first three decades of operation. Established
after a ‘second wave’ of longitudinal research that focused on the understanding of the etiology of chronic illness, which included the celebrated Framingham Heart Study (1948), the BLSA, like many other cohort studies of aging, lay at the intersection of two different understandings of the institutional configuration of expertise in the management of aging, health and illness. On one side, there was a set of actors who proposed to use serial observation of individuals’ physiological parameters, everyday activities and ‘attitudes’ in order to clarify processes relating to the onset of age-related illnesses such as heart disease or stroke, within a population. The focus was in establishing normative ‘risk’ parameters for specific age groups in relation to specific etiologies. On the other side of the argument, there was the ambition to establish new methodological conventions that would lead to being able to measure and chart ‘personalized age’ (biological, psychological, social). Actors on this side of the controversy argued that such individualized ranking measurements would instead support an alternative configuration of administrative and health care institutions, one where ‘older people’ would be treated as individuals and not merely as members of an age-group. This tension provides the key to understand the emergence of a critique of BA research, which will be the focus on the next section.

Unpacking biological age

In the previous section, I have analyzed how the search for BA became a key focus of gerontological research in the middle years of the 20th century. In this process, an emerging consensus was formed about the normative and political aims of developing individualized age measurements that would do away with the homogenizing effects of the widespread institutional use of CA. Closely linked to this was the collective articulation of the methodological procedures through which BA measurements should be developed. As I have suggested, such procedures challenged some of the established knowledge making conventions in epidemiology and statistics, a field that had done much to entrench CA as a tool in the modern administrative apparatus (see above). It did not take long for those methodological challenges to be reckoned with.

In 1977, the National Institute of Aging organized the 2nd Epidemiology of Aging Conference in order to “derive new knowledge to advance our understanding of the underlying causes of the aging process and help us separate disease from aging.”
(Butler 1980:4) This was an important occasion because it brought together researchers mostly used to working within laboratory settings – e.g. Richard Adelman, a biochemist from Temple University – with those concerned with studying aging, illness and health in community-dwelling populations, such as William Kannel, one of the investigators in the Framingham Heart Study.

The first session of the conference, chaired by Samuel Greenhouse - who as head of the Division of Statistical Methods at the US Public Health Service had previously been involved in the refinement of statistical methods in federally-funded studies such as the BLSA - was concerned with “attempt[ing] to formulate a definition of aging other than that of chronological age.” (Haynes and Ross 1980:vii) Two key presentations were included in this session. In the first, Dr. Adelman revised current experimental work on biological parameters of aging and concluded that epidemiological work on BA should go beyond observation, and record “altered physiological response to hormones, drugs, antigens, exercise and other conditions of stress.” (Adelman 1980:13)

The second key presentation of the conference was provided by Dr. Paul Costa and Dr. Robert McCrae, two psychologists working at the intramural research institute of the NIA in Baltimore. Entitled “Functional Age: Conceptual and Empirical Critique,” the presentation reviewed the existing studies focusing on individualized age measurement from Benjamin onwards, to conclude that they represented “a prescientific approach which showed considerable promise […] but which would be best abandoned.” (Costa and McCrae 1980:27) Their critique hinged on three main arguments. First, that it was illogical to want to devise a composite metric when it was known that different physiological systems age at different rates (see above). Second, that individuals’ rate of aging was unlikely to be constant through their life span. Third, that using multiple regression models to compose and validate the metric was methodologically flawed. In this last regard, Costa and McCrae argued that it was not only inappropriate to use the same battery for different age groups but also, and more importantly, that using CA as the criterion of validation of BA rendered the latter measurement somewhat useless, if the intention was to replace CA as a measure of age (see also Ingram 1988).
Because of these methodological problems, BA could not, in their view, replace CA as an individualized assessment of individuals’ capacity to work or of the effects of age-retarding interventions (see also Ludwig and Smoke, 1980). They concluded:

Chronological age is indeed a dummy variable, and research should attempt to replace it in every case with an account of the real etiology of age-related changes. The substitution of a new dummy variable, functional age, will not be a step forward on this process; indeed, it would be a step back (Costa and McCrae, 1980:45)

This attack was directed at the very heart of the socio-technical imaginary proposed for BA, i.e. that it would enable institutional and technological change underpinning older people’s health and socio-economic participation. Their suggestion was that it did not. Instead, they argued that the current CA-based system had many advantages. One key advantage related to the universalistic nature of CA. This made “chronological age […] completely democratic, possessed equally by rich and poor, man and woman, healthy and sick.” (Costa and McCrae 1980: 44; see also Kohli 1986). The invocation of the political ideal of equality upon which CA collection and calculation were based aligned Costa and McCrae’s critique with a defense of the modern administrative apparatus. Theirs was a ‘reformist’ approach to the inadequacies of CA. Linked to this approach, the second key advantage, referred to in the quote above, was that using CA it should be possible to build “for every case” profiles of “real etiology of age-related changes.” (emphasis added) In other words, CA was key for the implementation of personalized approaches in health care and social policy.

The significance of Costa and McCrae’s paper, evidenced by repeated printings of the paper and high number of citations, is that it turned the tables on the debate. Whereas proponents of BA had been supported by a consensus around the inadequacies of CA, a consensus in which Costa and McCrae included themselves, their paper questioned whether the solution proposed had any advantages in relation to the current CA-based system. As Greenhouse put it in the discussion following the presentation, “the onus [was now] on the other side.” (Greenhouse in Haynes et al 1980:47)

From the early 1980s onwards, starting with the 1981 Conference on Biological Markers of Aging referred to in the previous section, a series of concerted programs
of action – workshops, funding initiatives, consortia, bringing together an extended collective of researchers, clinicians, public and increasingly private funding agencies, mainly in the USA but also in Europe and Asia– were organized to build research capacity in this area. In this regard, they can be seen as addressing, sometimes explicitly, the methodological issues raised by the 1977 presentation (Interview with Richard Sprott: 10 December 2012). Amongst these attempts to move BA from a ‘pre-scientific’ to a scientific stage of development, the NIA Biomarkers of Aging Funding Initiative between 1988 and 1998 is perhaps the most important turning point.

Supported on the back of raising political interest on the effectiveness of age-retarding interventions (Interview with Richard Sprott: 10 December 2012), the program aimed to develop a panel of biomarkers of aging by assessing the effects of caloric restriction on genetically homogeneous strains of rats and mice. Deploying the conventions and instruments of laboratory physiological research, the program leaders hoped to find statistically significant differences between ad-hoc and restricted diets on particular measures. However, by 1999, it was clear that a panel of biomarkers had not been obtained (Sprott 1999). This was a momentous setback for BA proponents because the reasons behind its failure could be linked to the methodological critique enunciated by Costa, McCrae and others almost 20 year earlier.

Already at the start of the program, McClearn (1989) had warned that the research program had two main assumptions that might turn out to be problematic: that there was a relationship between physiological systems which could be captured in one single or composite measure, and that this measure could be used for both constant and variable rates of senescence across the lifespan. These, in effect, reiterated part of the Costa and McCrae 1977 critique. In addition, it became clear to interested, expert observers that the research was not statistically powered to be able to find significant differences (Interview with Richard Miller, 23 January 2013).

But the most poignant critique addressed the institutional, normative assumptions of the biomarkers of aging program. The fact that it was articulated by someone who just 10 years earlier could be considered to be in the BA camp, Richard Adelman, made it all the more piercing. Adelman was of the view that,
history will regard research related to the identification of such biomarkers of aging as no more than of fleeting significance. Sophisticated research entails so much more than pure empiricism.[…] efforts to characterize descriptively, as well as to solve the problems of old age, at least in my opinion, should not take precedence over the pursuit of new knowledge regarding the mechanisms of aging. […]Furthermore, it is a too often forgotten lesson of the history of science that the most significant societal advances usually result not by administrative design of the funding agencies, but instead, serendipitously from straightforward, non-targeted, investigator-initiated, high quality basic research. (Adelman 1987:227-29)

In understanding these remarks it is important to bear in mind that Adelman served as President of the Gerontological Society between 1986 and 1987. In this position, he had advocated, to funders and policy makers, a focused approach to the development of basic definitions of aging, supported by fundamental research within and across diverse disciplines (Achenbaum 1995:130-31). He was concerned that, since its establishment in 1974, the NIA had shifted its orientation from researching the mechanisms of aging to focusing on the diseases of aging, a process he later described as the ‘alzheimerization of aging’ (Adelman 1995; see also Moreira and Palladino 2009).

His view was that the biomarkers program partook in these changes where policy makers attempt to direct research towards societal and economic impact, often with little understanding of what he saw as the role of serendipity in the ‘history of science’. In this invocation of the “history of science,” and of the value of curiosity driven research, Adelman also set forth a vision of division of labor within biomedical sciences, whereby basic research is institutionally separated from the development of technological innovation and the application of research in the clinic by qualified physicians. Blurring of these section of the research endeavor would lead, in his view, to unnecessary affective burdens on the activities of scientists and a resulting lack of focus on “the pursuit of new knowledge regarding the mechanisms of aging.” (See Adelman 1987: 227) The socio-economic and biomedical imperatives for developing a measure of BA did not thus justify the implementation of ‘stop-gap measures’ (Adelman 1987:228) in aging research, particularly if the pragmatic weaknesses of the BA approach, identified by Costa and McCrae, were to be considered.
The years following the end of the NIA program are acknowledged to have been difficult for BA proponents (Butler et al. 2004; Interview with Richard Miller, 23 January 2013), having lost support not only from previous political allies and research policy makers but also from within their own side. Responding to this, BA proponents shifted their attention from federal funding to gathering support and funding from independent, privately funded research institutes and foundations, such as the International Longevity Center, the Ellison Medical Foundation or the American Foundation for Aging Research. Another important strategy appears to have been to make the critique of BA as a point of departure for new research programs. Indeed, in an interview with cell biologist Thomas von Zglinicki (20 June 2013), much emphasis was put on ensuring that possible BA measurements were robustly tested against the claims of critics before publicizing any advancement in research in this field. In this, BA proponents’ strategy appeared to revolve around honing the public ‘presentation of the scientific self’ (Hilgartner 2000)

Just two years after the failed NIA program ended, Dr. Richard Miller, a molecular biologist at the University of Michigan, argued that,

The anti-biomarker camp is heartened by deconstructions (Costa and McCrae, 1980) of previous attempts to develop mathematical composites of age-sensitive traits that could be mistaken for quantifiable indices of "generalized aging." The critics of linear regression and principal components approaches to biomarker construction argue, I think convincingly, that evidence for some correlation among different age-sensitive traits does not constitute proof that these traits are suitable indices for some unmeasured (and perhaps unmeasurable) "rate of aging." Previous attempts to develop batteries of biomarker assays have also stumbled over statistical obstacles, including failure to control for age-independent differences among subjects in the traits of interest [But] in my view, though, these critics jump too quickly from "has not been done" to "cannot be done." At the heart of their argument is the allegation that biomarkers of aging cannot be developed because "there is no single rate of aging" […] These critics may be right, but the question is too important to be decided by edict rather than by experimentation. (Miller 2001:21)

Acknowledging the methodological and statistical difficulties encountered by BA research in the past, Miller repositions Costa and McCrae critique as a
‘deconstruction’ underpinned by the belief that “there is no single rate of aging.” Miller then contrasts this ‘allegation’ with a scientific attitude of skepticism, openness to evidence and an experimental approach. The expert acknowledgment of the challenges faced by BA research also connotes modesty and reinforces Miller’s aims to portray the dispute as impersonal. In drawing on such established normative ideal of scientific conduct (Shapin 2008), Miller was also denouncing the irrational, interested ground upon which the institutional apparatus defended by Costa and McCrae lay, which could only be established by ‘edict’ and the power of authority. In this denunciation, it was not insignificant the fact that Costa and McCrae’s original research had been partially funded by the Council for Tobacco Research USA. In many respects, Miller’s can be seen as a direct response to the ‘anti-biomarker camp’s’ defense of CA and its democratic, egalitarian nature. For BA proponents, such defense amounted to an undermining of the foundations of an open, pluralistic society.

Following this line of normative justification for BA research, Miller and his supporters came to articulate a more consistent critique of what they saw as the political obstacles to their program. This led them to a critical analysis of biomedicine, which they saw as eminently reliant on CA (Kirkwood 2001). However, instead of, like Adelman, condemning the incursion of aging research into the biomedical model of R&D, BA proponents highlighted the failures of such a model, and began constructing promissory scenarios linked to the returns from investigation into age-retarding processes and agents (Miller 2009).

Against Costa and McCrae’s proposal that the focus should be to understand “for every case […] the real etiology of age-related changes,” proponents of BA pointed to the weak returns in terms of life-years gained from a specific etiology approach, in fields like cancer or heart disease. This line of attack on biomedicine institutional apparatus was fruitful because it resonated against a backdrop of other condemnations of biomedicine for its decreasing rate of therapeutic innovation (e.g. Fanu 1999). Importantly, it opened the ground for aligning new actors in a renewed approach to aging, health and illness, such as the Wellcome Trust (2006) or the European Commission (MARK AGE Program 2008-13).

As molecular biologist Robin Holliday put it in an interview to the journal Biogerontology,
That is the way most doctors look at disease. Whenever there is any particular disease, they tend to treat that specific disease, and ignore the fact that other parts of the body may be failing. […] In my opinion there needs to be a whole re-appraisal of the field of age-related disease, and much more emphasis on aging research. What is important is to look for the origin of each disease and then try to prevent it from happening. (Holliday in Rattan 2002:34)

As Paolo Palladino and I (2009) have argued, this critique of biomedicine strategically positioned aging researchers - and BA proponents – apart both from rejuvenation activists and ‘futurologists’ (Turner 2009) and from the ‘disease-mongering’ practices of the medico-industrial complex (Moynihan, Heath and Henry 2002). Instead of promising to cure aging or disease, BA proponents portrayed themselves as part of a scientific focus on aging research that would bring about preventative, health enhancing interventions and technologies (Butler et al 2008). Instead of mending particular organ systems, BA proponents aimed at understanding and measuring the common ‘origin’ of all age-related disease.

This meant questioning and challenging the authority of the medical profession. Here again, they were not isolated. Indeed, since at least the 1970s, in the US and Europe, the medical profession had faced what Light (1988) described as a ‘buyers’ revolt’ in the form of increased consumerism, managerialism, corporatization and standardization. The latter, most commonly associated with the evidence-based movement, was particularly significant because it was linked to shifts in the knowledge base of medical practice (Timmermans and Kolker 2004) and an emphasis on an ‘experimental ideal’, firmly linking procedures of knowledge making with the governance of pluralistic societies (Moreira 2012:87-110). This helps understand not only Miller’s response to Costa and McCrae, but also BA proponents’ renewed confidence in a standard-driven process of institutional transformation. BA became thus the central actor in what actor-network theorists would label as the ‘program of action’ of contemporary bio-gerontology.

But this was also a potential weakness, as the transformation of biomedicine could only be actualized if the elusive age measurement was found. This is recognized by the ‘anti-biomarker camp’, and no significant attempts have been made to redress the balance in the controversy in the last decade or so (Robert McCrae, personal communication). For them, the ‘onus is [still] on the other side’. Because no
biomarker of aging or battery of tests “have so far proven [to be] a true measure of the underlying aging process,” (American Federation of Aging Research 2011) it is still the case that it is possible to deliver individualized care if the “physician [is able] to assess the status of the kidneys, heart, and liver, and to evaluate them relative to age, sex, and other norms.” (Costa and McRae 1988:213) For opponents of BA, standards or norms of health and function are there to assist physicians in their diagnostic and prognostic assessments, providing pastoral guidance for the patient in front of him/her. This can be done with and through the existing institutions of biomedicine.

Presently, the controversy appears to have reached a stalemate, with each ‘camp’ entrenched in their respective institutional configurations. As a consequence, both CA-derived niche risk assessments and attempts to measure BA are shrouded in uncertainty, and both versions of ‘personalization’ remain unstable. This in turn has consequences for the implementation of approaches to manage aging and its relation to health and illness. Indeed, it could be said that it impacts on how aging is understood and enacted within the biosciences, remaining an ‘unsolved problem’ more than 60 years since Peter Medawar’s famous lecture on the subject (Medawar 1952). As Sprott admits as recently as 2010,

While there is no definition of aging that is universally accepted, most gerontologists at least do no object too strenuously to the following: “Aging . . .is the progressive deterioration of virtually every bodily function over time.” (Austad, 1997) This definition and others implies that aging is the result of more than one on-going process and that some systems may be (probably are) affected independently. How much of the observed deterioration is due to basic processes and how much is due to disease processes is the subject of continuous discussion. (Sprott 2010:1)

Conclusion

In concluding, I would like to return to the telomere controversy described in the beginning of the paper. The conflict between Blackburn and Greider hinged on whether telomere length could be used as a biomarker of aging in humans, with Blackburn appearing to be confident it could, putting her scientific reputation behind a commercial company providing BA measurement. Greider, on the other hand, remained skeptical about the evidence-base supporting the clinical use of this bio-
technology. Whereas, at the start of the paper it would be reasonable to see Greider as an opponent of the biomarker approach, by now, drawing on the analysis presented in this paper, it should be clear that she should be more accurately characterized as being in the ‘biomarker camp’. Likewise, while Blackburn was described as “selling a biological age test,” (Dickinson 2011) her approach is in reality more in line with critics of BA, advocating that individualization of care can be achieved through health assessment and risk profiling for specific illnesses, with the assistance of a physician (see above).

As I have suggested in the introduction, both the telomere controversy and the wider debate on BA measurement has hinged on a divergence on how to design standards to implement individualized approaches to the management of aging, health and illness. In this respect, the BA controversy is best understood as part of a wider shift in standardization in late modernity, whereby measures and scales promise individualization and ‘personalization’ of technologies or services. As Busch (2011) and Epstein (2009) have suggested, the goal is neither to develop a universalistic standard, applicable to all, nor to rely solely on cases’ uniqueness. Instead, in domains as diverse as transport, communication technologies, biomedicine or aging, experts focus on the conception, validation and implementation of standards which purport to identify the combination of unique characteristics of persons that are relevant to a specific market, service, technology, type of work, etc. The shift in standardization practices alone cannot however explain the emergence of an interest in BA measurement, as this is distinctively linked to the dynamics of the sciences of growth, development and senescence during the 20th century.

As I suggested in the section Personalizing age measurement, the concept of BA surfaces at the intersection between a political program aiming to reform modern education, work and welfare institutions and a specific approach to knowledge making that focus on trajectories of individuals through key stages in their life. This supported a consistent denunciation of the nefarious effects of using CA as a metric to distribute rights and duties in the modern society. From the 1930s onwards, such critique began to highlight the irrationalities and unfairness that arises from the disconnect between CA-based systems and the diversity of ‘capacities, needs and desires ‘of older people. In the three decades after WWII, this critique also backed the
emergence of a consensus around the normative purpose of BA and the best methods used to measure it.

There were however contradictions in using methods that rely on regression to the mean to develop individualizing standards. This became increasingly clear when negotiating the methodological approaches to be used in longitudinal studies of aging, and it is of no surprise that the key critical assessment of BA was voiced within a forum on the epidemiology of aging. Costa and McCrae’s critique not only clearly identified the conceptual and methodological weakness of BA, but also contended that CA was a benevolent, intrinsically egalitarian standard which could be used in reforming health and welfare institutions. Their critique helped also to define the ‘matter of concern’ (Latour 2005:87-120), and to structure the duality of views of, and normative approaches and epistemic commitments to the relationship between aging, health and illness. Particularly after the failure of the NIA Biomarkers of Aging program, this diversity became divided into what those involved in the controversy designated as ‘camps’, a term which connotes the level of uncertainty and conflict about aging measurement in contemporary bio-clinical research on aging.

As I mentioned above, one of the key questions concerning social scientists interested in aging and the life course has been to understand whether, how and why transitions to adulthood and old age have become de-institutionalized, or less structured by normative obligations. As health and the ‘obligation to stay active’ increasingly bear on how social and political institutions delineate “the trade-offs across the lifespan,” (Daniels 2009:39; also Lassen and Moreira 2014) it becomes important to understand how current framings of the aging process might shape the range of possible engagements with life course processes available to individuals. This paper shows that there are weighty uncertainties not only about the framing but crucially about the infrastructure – the standards – supporting individuals and organizations’ decisions about life course trajectories. I have argued that these uncertainties derive from mutually reinforcing, divergent ways of enacting the relationship between normative ideals of the life-course and methodological approaches to knowledge making to understand and manage the relationship between aging, health and illness. These divergences concern not only CA or BA themselves but more importantly how to organize institutions to measure, monitor and manage age in contemporary societies. Such diagnosis of the current infrastructural apparatus of classification and
quantification of age should be taken in consideration when pursuing further research on the life course transitions as institutionally, normatively embedded action.

Archives

Alex Comfort Papers: Special Collections Archive at the University College London Library. References to this resource coded as AC: [box number].

Nathan W. Shock Collection: Bentley Historical Library at the University of Michigan. References to this resource coded as NWS: [box number].

References


Bryson, Dean. 1998.” Lawrence Frank, knowledge and the production of the social.” Poetics Today, 19, 401-421


Connor, Steven. 2012. “Want to know how old you really are? Then join the queue.” The Independent, 27 December 2012


Hughes, Mary Elizabeth and Linda J. Waite. 2007. The aging of the second demographic transition. New York: Springer.


Miller, Richard A. 2009. “Dividends” from research on aging—can biogerontologists, at long last, find something useful to do?.” The Journals of Gerontology Series A: Biological Sciences and Medical Sciences, 64(2), 157-160.


