The reported healthy ageing gene expression score: lack of predictive value in two cohorts

A response to: **A novel multi-tissue RNA diagnostic of healthy ageing relates to cognitive health status** by Sood S, Gallagher IJ, Lunnon K, Rullman E, Keohane A, Crossland H, et al. *Genome Biology*; 2015;**16**:185.

Luke C. Pilling[1], Lorna W Harries[2], Dena G Hernandez[3], Andrew B Singleton[3], George A. Kuchel[4], Luigi Ferrucci[5], David Melzer[1]

Affiliations:

1. Epidemiology and Public Health Group, Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, United Kingdom.
2. RNA Mechanisms of Disease Group, Institute of Biomedical and Clinical Science, University of Exeter Medical School, Exeter, United Kingdom.
3. Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, MD, United States.
4. Division of Geriatrics, University of Connecticut Health Center, Farmington, Connecticut
5. Clinical Research Branch, National Institute on Aging, Baltimore, MD, United States.

Corresponding author:

Professor David Melzer,

Epidemiology and Public Health,

University of Exeter Medical School,

Medical Research, RILD level 3,

RD&E Wonford, Barrack Road, Exeter, EX2 5DW, UK.

E-mail: D.Melzer@exeter.ac.uk

Phone: +44 (0) 1392 406751

Abstract

Sood *et al.* report a multi-tissue RNA signature "predictive of human health, using only peripheral blood samples". We tested this score in blood in two independent, larger cohorts and found no associations with age or related phenotypes, including strength, interleukin-6 or mortality.

Correspondence

Sood *et al.* (1) reported identifying a "150-gene healthy ageing classifier" of expression markers which accurately (>90%) classified muscle tissue as older (65-year-old, n=15) or younger (25-year-old, n=15) (1). This age classifier was then tested in publicly available muscle gene expression databases and in samples from skin, blood and brain tissue, and is reported to have performed well across data from various expression arrays. Using data from 108 participant's muscle samples in the ULSAM cohort (all about 70 years of age and male), the authors collapsed the classifier into a single "healthy ageing gene score" HAGS (methods described in detail in (1)). The HAGS at age 70 was positively associated with creatinine clearance 12 years later (p=0.009) and negatively associated with 20-year survival (p=0.03, 19 deaths in 108 participants). HAGS was not cross-sectionally associated with ageing associated phenotypes including blood pressure, cholesterol, 2-hour glucose tolerance test, or renal function at baseline in ULSAM. Sood *et al.* report that applying the HAGS to blood samples accurately (up to 86%) identified individuals with Alzheimer's from controls (2 cohorts, total n=224).

Here we sought to further validate the HAGS using whole blood-derived expression data from two well-characterized human cohorts including wide age ranges and both sexes; the Italian aging cohort, InCHIANTI (2), and the Baltimore Longitudinal Study of Aging, BLSA (3). We hypothesized that a valid gene expression ageing score should be positively correlated with chronological age and should predict mortality. Advancing age on the gene score should be also associated with ageing related phenotypes including declining muscle strength, cognitive and renal function, plus increasing inflammation. In addition and perhaps most crucially, a valid ageing score should be a better predictor of these outcomes than chronological age (e.g. be significantly associated with mortality after adjustment for chronological age).

Testing the validity of the HAGS in two independent cohorts

Sood *et al.* report successfully applying the 150 Affymetrix array derived gene set in data from Illumina Human HT-12 v3 microarrays used in skin and blood samples, with 128 genes included. In our Illumina HT-12 based data, we found that 119 genes had a corresponding probe in the annotation for the Illumina Human HT-12 v3 array, of which 77 genes were significantly expressed above background (v3 array, $p<0.01$ in at least 5% of participants (4)) in InCHIANTI and 66 (v4 array) in BLSA. We therefore calculated associations with both the genes expressed above background and the full n=119 sets separately. We replicated the methods of Sood *et al.* which state that a participant with the highest expression of a gene that decreases with advancing age is ranked as 1, so it follows that the "youngest" profile has the smallest HAGS value.

Details of the sample collection, quality assurance and analysis of the InCHIANTI data are set out in Harries *et al*. 2011 (4), and similar methods were used in BLSA. HAGS had an approximately Gaussian distribution in both cohorts and we therefore used HAGS as the dependent variable in robust linear regression models (STATA v13). In each case the analysis was adjusted for age, sex, microarray batch and basic white blood cell subtype counts (number of lymphocytes, monocytes, eosinophils and basophils). InCHIANTI was also adjusted for each of two study sites and BLSA was adjusted for race (white/other). For the association with all-cause mortality, Cox proportional hazards models are used with the same covariates as above.

Summary statistics for the study participants are provided in **Table 1**. Maximum follow-up time periods for InCHIANTI and BLSA were approximately 6.7 years and 6 years, respectively.

Table 1: Cohort summary statistics

	InCHIANTI	**BLSA**
N	695	763
Age (mean, SD)	72 (15.3)	68 (12.9)
Age (min, max)	30 - 104	31 - 98
Females (n, %)	382 (55%)	368 (48%)
Race, white (n, %)	695 (100%)	517 (68%)
Died during follow-up (n, %)	156 (23%)	45 (6%)
Years to death since RNA visit (mean, SD)	3.45 (1.77)	3.04 (1.48)
Years to death since RNA visit (min, max)	0.23 - 6.66	0.25 - 6.02
Healthy Ageing Gene Score (mean, SD)	351 (65)	394 (187)
Healthy Ageing Gene Score (min, max)	117 - 549	7 - 763

We tested associations with the phenotypes listed in **Table 2**. HAGS based on genes expressed above background was not associated with our studied ageing phenotypes in cross-sectional analyses in either cohort. Renal function, estimated by Cockcroft-Gault creatinine clearance, was significantly but positively associated with HAGS in InCHIANTI i.e. "older" gene scores were paradoxically associated with increased or better creatinine clearance, with a similar but non-significant effect size in BLSA. Cox survival models showed no association between HAGS and mortality over 6-years (**Figure 1**), either in a categorical analysis (**Table 2**) or in a continuous analysis of the trend (HR=0.99, p=0.84, and HR=0.99, p=0.55, in InCHIANTI and BLSA respectively). Data were available on Mini Mental State Examination score (MMSE, a widely used cognitive impairment measure) in the InCHIANTI cohort, but we found no associations with HAGS using linear or categorical analyses (linear model: coef= -8.15, p=0.266).

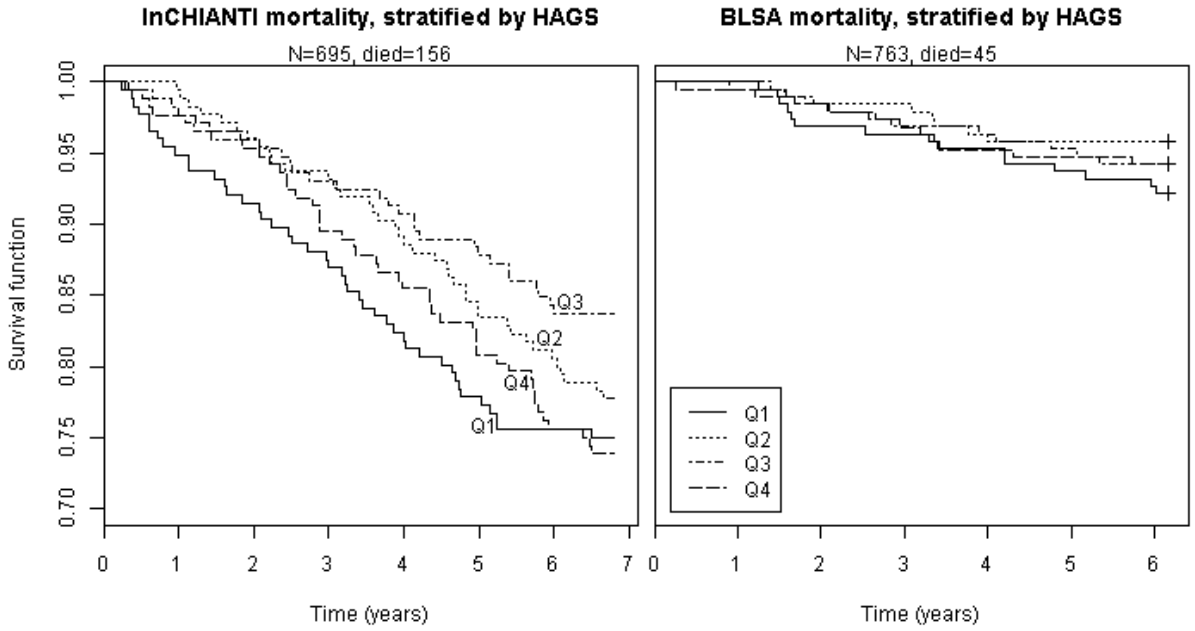
Table 2: Associations between HAGS and "ageing" phenotypes

	InCHIANTI				***BLSA***			
Linear regression (robust)	**Coefficient**	**95% CIs**		**p-val**	**Coefficient**	**95% CIs**		**p-val**
Age (years)	-0.138	-0.43	0.15	0.348	0.219	-0.50	0.94	0.548
BMI (kg/m2)	0.700	-0.38	1.78	0.205	1.549	-0.38	3.47	0.115
Smoking status (ever/never)	3.291	-6.06	12.64	0.490	1.053	-16.87	18.97	0.908
Max hand-grip strength (kg)	-0.235	-0.92	0.45	0.501	1.026	-0.27	2.32	0.120
Creatinine clearance (mL/min)	0.328	0.09	0.57	0.008	0.490	-0.03	1.01	0.065
Interleukin-6 serum (pg/mL) ¥	0.553	-6.75	7.85	0.882	-18.803	-49.62	12.01	0.231
Mortality (Cox PH model)	**HR (95% CIs)**	**p-val**	**n / N (%) ***		**HR (95% CIs)**	**p-val**	**n / N (%) ***	
Quartile 4 (reference)	1.000	-	45		1.000	-	11	
	-		172	(26%)			190	(6%)
Quartile 3	0.778	0.338	28		0.509	0.201	11	
	(0.47 to 1.30)		172	(16%)	(0.18 to 1.43)		191	(6%)
Quartile 2	0.818	0.417	39		0.365	0.089	8	
	(0.50 to 1.33)		175	(22%)	(0.11 to 1.17)		191	(4%)
Quartile 1	1.100	0.723	44		0.869	0.819	15	
	(0.65 to 1.86)		176	(25%)	(0.26 to 2.88)		191	(8%)

¥ Interleukin-6 values log-transformed prior to analysis, and outlier values >4 SD from the mean are excluded

* n = number of deaths during follow-up, N = total participants in this subset, % = n/N*100

Figure 1: Kaplan-Meier survival plots show the relationship between HAGS and mortality status



Plots of proportion surviving by year of follow-up and quartile of HAGS in InCHIANTI and BLSA separately. Q1 are the lowest (youngest) quartile of HAGS scores. Figure generated in R (version 3.2) using package `survival` (v2.38).

We reran the aforementioned analyses including all available Illumina HT12 v3 HAGS gene probes (i.e. not just the probes significantly expressed above background levels), but the results were not significantly different.

We also saw no significant differences in associations when only including male participants in the analysis. Finally, Sood *et al.* do not report adjusting for age as their cohorts were all age-matched; the only difference in our results when not including age as a covariate (or cell counts, which are correlated with age) was that creatinine clearance was no longer associated with HAGS (coef=0.086, p=0.31).

Healthy Aging Gene Score in blood was not associated with aging traits

We have shown that expression of the reported health ageing gene score (HAGS) was not associated with chronological age or a battery of age-related phenotypes in blood samples from 2 independent human studies, each larger than the cohorts examined by Sood *et al*, except for a positive association between HAGS and creatinine clearance. As noted, Sood et al's methods state that a participant with the highest expression of a gene that decreases with age is ranked as 1, so it follows that the "youngest" profile has the smallest HAGS value. Our observation (also made by Sood *et al.*) that higher HAGS is associated with higher creatinine clearance therefore appears paradoxical. Moreover, we did not observe a significant relationship between HAGS and risk of mortality over 6 years in two relatively large independent cohorts, not even in the direction presented in Sood et al, which shows lower HAGS scores having higher mortality (Supplementary figures S1A and S1B).

Although the HAGS appears not to be predictive of ageing phenotypes in blood, there is accumulating evidence for robust changes in gene expression in blood associated with age and related phenotypes. In 2013 we showed that large numbers of gene expression markers were strongly associated with age in blood in the InCHIANTI study (5), and that a "biological age" estimation with 5 selected genes was associated with chronological age, muscle strength, albumin, IL6 and renal function (blood urea nitrogen) in the appropriate directions. Using the 6.7 years of follow-up data we also found that participants predicted to be >10 years 'biologically' younger than their chronological age had substantially lower mortality (Hazard Ratio=0.52 95%CI 0.31 to 0.88) compared to participants predicted >10 years older than their chronological age.

In a recent meta-analysis of nearly 15,000 participants from 14 cohorts, Peters *et al.* report that expression of 1,497 genes in human whole blood samples are robustly associated with age, with 25% of these also associated with age in brain tissue (6). The meta-analysis provided good replication of the earlier results from InCHIANTI alone (4,5). However, of the 150 genes identified by Sood *et al.* only 10 genes were also found to be associated with age in this large-scale meta-analysis, suggesting that 93% of expression markers in HAGS are unlikely to be associated with chronological age in human blood.

We have shown in two relatively large independent cohorts that the recently reported "healthy aging gene score", measured in blood, was not associated with chronological age or several age-related phenotypes and was not predictive of mortality over 6 years. Valid blood based gene expression markers of age have been previously reported, and predictors based on them perform well against chronological age in valid assessments.

Declarations

All authors have read the manuscript and provided input, and declare no competing interests.

References

1. Sood S, Gallagher IJ, Lunnon K, Rullman E, Keohane A, Crossland H, et al. A novel multi-tissue RNA diagnostic of healthy ageing relates to cognitive health status. Genome Biol. Genome Biology; 2015;16:185.
2. Ferrucci L, Bandinelli S, Benvenuti E, Di Iorio A, Macchi C, Harris TB, et al. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. J Am Geriatr Soc. 2000 Dec;48(12):1618–25.
3. Shock NW, Greulich RC, Costa, Paul T J, Andres R, Lakatta EG, Arenberg D, et al. Normal Human Aging: The Baltimore Longitudinal Study on Aging. NIH Publ Washington, DC. 1984;
4. Harries LW, Hernandez D, Henley W, Wood AR, Holly AC, Bradley-Smith RM, et al. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging Cell. 2011 Oct 13;10(5):868–78.
5. Holly AC, Melzer D, Pilling LC, Henley W, Hernandez DG, Singleton AB, et al. Towards a gene expression biomarker set for human biological age. Aging Cell. 2013 Jan 12;12(2):324–6.
6. Peters MJ, Joehanes R, Pilling LC, Schurmann C, Conneely KN, Powell J, et al. The transcriptional landscape of age in human peripheral blood. Nat Commun. 2015 Oct 22;6:8570.