## 1 Extreme MetaboHealth scores in<br>
2 three cohort studies associate with<br>
3 plasma protein markers for  $_2$  three cohort studies associate with<br>  $_3$  plasma protein markers for<br>  $_4$  inflammation and cholesterol 3 **plasma protein markers for**<br>4 **inflammation and cholesterol**<br>5 **transport.**<br>6 **D. Bizzarri<sup>123</sup>. E.B. van den Akker<sup>123</sup>. M.J.T. Reinders<sup>2,3</sup>. R. Pool<sup>4,5</sup>. M. Beekman<sup>1</sup>.** 4 **inflammation** and cholesterol<br> **17 inflammation** and **cholesterol**<br> **17 EXECUTE:**<br> **1 EXECUTE:**<br> **1**

5 **transport.**<br>
6 D. Bizzarri<sup>1,2,3</sup>,<br>
7 N. Lakenberg<sup>1</sup>, N.<br>
8 M.A. Ikram<sup>10</sup>, M.<br>
9 P.E. Slagboom<sup>1,14#</sup> D. Bizzarri<sup>4,2,3</sup>, E.B. van den Akker<sup>4,2,3</sup>, M.J.T. Reinders<sup>2,3</sup>, R. Pool<sup>4,3</sup>, M. Beekman<sup>4</sup>,<br>
N. Lakenberg<sup>1</sup>, N. Drouin<sup>6,7,8</sup>, K.E. Stecker<sup>7,8</sup>, A.J.R. Heck<sup>7,8</sup>, E.F. Knol<sup>9</sup>, J.M. Vergeer<sup>10</sup>,<br>
M.A. Ikram<sup>10</sup>, M N. Lakenberg<sup>+</sup>, N. Drouin<sup>9,7,8</sup>, K.E. Stecker<sup>7,8</sup>, A.J.R. Heck<sup>7,8</sup>, E.F. Knol<sup>-7</sup>, J.M. Vergeer<sup>26</sup>,<br>
M.A. Ikram<sup>10</sup>, M. Ghanbari<sup>10</sup>, A.J. van Gool<sup>11</sup>, BBMRI-NL<sup>12</sup>, D.I. Boomsma<sup>4,5,13,14</sup>, and<br>
P.E. Slagboom<sup>1,14#</sup>

M.A. Ikram<sup>40</sup>, M. Ghanbari<sup>20</sup>, A.J. van Gool<sup>14</sup>, BBMRI-NL<sup>42</sup>, D.I. Boomsma<sup>4,5,13,14</sup>, and<br>
9<br>
P.E. Slagboom<sup>1,14#</sup><br>
<sup>1</sup> Department of Biomedical Data Sciences, Molecular Epidemiology, LUMC, Leiden, The Netherlands<br>
<sup>2</sup> P.E. Slagboom<sup>+++++</sup><br>
10 <sup>1</sup> Department of Biomedical Data Sciences, Molecular Epidemiology, LUMC, Leiden, The Netherlands<br>
<sup>2</sup> Department of Biomedical Data Sciences, Leiden Computational Biology Center, LUMC, Leiden, The

<sup>3</sup> Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands

Ĭ

<sup>2</sup> 11 Department of Biomedical Data Sciences, Leiden Computational Biology Center, LUMC, Leiden, The<br>
21 Netherlands<br>
<sup>3</sup> Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands<br>
<sup>4</sup> Department of 12<br>
14<br>
15<br>
<sup>3</sup> Delft Bioinfo<br>
<sup>4</sup> Department<br>
<sup>5</sup> Amsterdam<br>
<sup>5</sup> Amsterdam<br>
<sup>7</sup> Biomolecu<br>
<sup>7</sup> Biomolecu<br>
<sup>7</sup> Biomolecu<br>
<sup>7</sup> Biomolecu<br>
<sup>8</sup> Netherland<br>
<sup>8</sup> Netherland<br>
<sup>8</sup> Netherland <sup>3</sup> Delft Bioinformatics Lab, Delft University of Technology, Delft, The Netherlands<br>
<sup>4</sup> Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands<br>
<sup>5</sup> Amsterdam Public Health Research Examples and Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands.<br>
15 <sup>5</sup> Amsterdam Public Health Research Institute, Amsterdam, The Netherlands<br>
<sup>6</sup> Leiden Academic Centre for Dru Famsterdam Public Health Research Institute, Amsterdam, The Netherlands<br>
<sup>6</sup> Leiden Academic Centre for Drug Research, Leiden University, Leiden, Neth<br>
<sup>7</sup> Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for The Mathematical Centre for Dug Research, Leiden Onversity, Leiden, Netherlands<br>
16 Ultrecht Institute for Pharmaceutical Sciences, Utrecht University, Utrecht, The Netherlands<br>
<sup>8</sup> Netherlands Proteomics Center, Utrecht, <sup>7</sup> Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and

<sup>8</sup> Netherlands Proteomics Center, Utrecht, The Netherlands

<sup>10</sup> Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands

17 Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and<br>
17 Biomolecular Mass Spectrometry and Proteomics, Bijvoet Center for Biomolecular Research and<br>
17 Biomolecular Research and<br> <sup>2</sup> Utrecht Institute for Frammaceutical Sciences, Ortecht Inversity, Ortecht, The Netherlands<br>
<sup>3</sup> Netherlands Proteomics Center, Utrecht, The Netherlands<br>
<sup>9</sup> Center of Translational Immunology and Dermatology/Allergolog From Schemes Protein, Scheme, The Netherlands<br>
<sup>9</sup> Center of Translational Immunology and Dermatology/Allergology, University Medical Center Utrecht,<br>
22 <sup>10</sup> Department of Epidemiology, Erasmus MC, Rotterdam, the Netherla 22 <sup>10</sup> Department of Epidemi<br>
<sup>10</sup> Department of Epidemi<br>
<sup>11</sup> Translational Metabol<br>
25 Center, Nijmegen, the Netherlands<br>
<sup>12</sup> BBMRI-NL: https://www.<br>
27 <sup>13</sup> Amsterdam Reprodu<br>
28 Netherlands. <sup>11</sup> Translational Metabolic Laboratory, Department of Laboratory Medical<br>
25 Center, Nijmegen, the Netherlands<br>
<sup>12</sup> BBMRI-NL: https://www.bbmri.nl.; see Consortium Banner Supplement S<br>
<sup>13</sup> Amsterdam Reproduction and Dev

<sup>12</sup> BBMRI-NL: https://www.bbmri.nl; see Consortium Banner Supplement S1

24 Medical<br>
25 Center, Nijmegen, the Netherlands<br>
<sup>12</sup> BBMRI-NL: https://www.bbmri.nl; see Consortium Banner Supplement S1<br>
<sup>13</sup> Amsterdam Reproduction and Development (AR&D) Research Institute, Amsterdam, The<br>
28 Netherla 25<br>25<br>26<br>29<br>29<br>30<br>31 26 Center, Nijmegen, the Netherlands<br>
<sup>12</sup> BBMRI-NL: https://www.bbmri.nl<br>
27 <sup>13</sup> Amsterdam Reproduction an<br>
29 <sup>14</sup> Dept of Complex Trait Genetics, C<br>
30 Universiteit Amsterdam<br>
<sup>15</sup> Max Planck Institute for the Biolo

27 BBMRI-NL: https://www.bomm.nl; see Consortium Banner Supplement S1<br>28 Netherlands.<br><sup>14</sup> Dept of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam, Vrije<br>30 Universiteit Amsterdam<br><sup>15</sup> Ma <sup>14</sup> Dept of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam, Vrije<br>
30<br>
<sup>15</sup> Max Planck Institute for the Biology of Ageing, Cologne, Germany<br>
32

31 Universiteit Amsterdam<br>
<sup>15</sup> Max Planck Institute fo<br>
32 15 1 Max Planck Institute for the Biology of Ageing, Cologne, Germany<br>3 2<br>The Max Planck Institute for the Biology of Ageing, Cologne, Germany<br>3 2

**NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.**

33 ABSTRACT<br>34 The MetaboHe<br>35 contains only a sr The MetaboHealth score is a highly informative health indicator in ageing studies and yet<br>contains only a small number of metabolites. Here we estimate the heritability of the score<br>in 726 monozygotic (MZ) and 450 dizygoti 36 in 726 monozygotic (MZ) and 450 dizygotic (DZ) twin pairs, and test for association with<br>37 plasma proteins by comparing extreme scoring individuals selected from two large plasma proteins by comparing extreme scoring individuals selected from two large<br>38 population cohorts -the Leiden Longevity Study (LLS) and the Rotterdam Study (RS) and 38 population cohorts -the Leiden Longevity Study (LLS) and the Rotterdam Study (RS) and<br>39 discordant monozygotic twin pairs from the Netherlands Twin Register (NTR).

40 The heritability for the MetaboHealth score was estimated at 40%. In 50 high and 50 low 39 1999 and the heritability for the MetaboHealth score was estimated at 40%. In 50 h<br>39 discordant monoster (NTR). The Scoring MetaboHealth groups from LLS and RS, we uncovered significant 41 scoring MetaboHealth groups from LLS and RS, we uncovered significant differences in<br>42 plasma proteins, notably in 3 (out of 15) cytokines (GDF15, IL6, and MIG), and 106 proteins 42 plasma proteins, notably in 3 (out of 15) cytokines (GDF15, IL6, and MIG), and 106 proteins<br>43 (out of 289) as determined by Mass Spectrometry based proteomics analysis. A high 43 (out of 289) as determined by Mass Spectrometry based proteomics analysis. A high<br>44 MetaboHealth score associated with an increased level for 42 serum proteins, 44 MetaboHealth score associated with an increased level for 42 serum proteins,<br>45 predominantly linked to inflammation and immune response, including CRP and HPT. A low 45 predominantly linked to inflammation and immune response, including CRP and HPT. A low<br>46 Score associated with decreased levels of 71 proteins enriched in high-density lipoprotein 45 predominantly linked to inflammation and immune response, including CRP and HPT. A low<br>46 score associated with decreased levels of 71 proteins enriched in high-density lipoprotein<br>47 (HDL) remodeling and cholesterol tr 47 (HDL) remodeling and cholesterol transport pathways, featuring proteins such as APOA1,<br>48 APOA2, APOA4, and TETN.

49 In MZ twins selected for maximal discordance within a pair we found 68 serum proteins 49 In MZ twins selected fo<br>50 associated with the Metal 49 In M2 Thuns selected for maximum discordance them a pair we found 68 servements associations<br>51 In observed in LLS and RS is likely explained by genetic influences. Taken together, our study 51 observed in LLS and RS is likely explained by genetic influences. Taken together, our study<br>52 sheds light on the intricate interplay between MetaboHealth, plasma proteins, cytokines, 52 sheds light on the intricate interplay between MetaboHealth, plasma proteins, cytokines,<br>53 and genetic influences, paving the way for future investigations aimed at optimizing this 53 and genetic influences, paving the way for future investigations aimed at optimizing this<br>54 mortality risk indicator. 54 and generic influences, paring the way for future investigation and at optimizing the<br>54 mortality risk indicator. 54 mortality risk indicator.

55 **INTRODUCTION**<br>56 As the global human<br>57 expected resilience of old 57 expected resilience of older individuals to support prevention and well-informed treatment<br>58 aimed at enhancing well-being [1]. Efficient disease prevention hinges on possibilities for 58 aimed at enhancing well-being [1]. Efficient disease prevention hinges on possibilities for<br>59 evaluating not only an individual's disease risk, but also the overall physiological vulnerability Example an enhancing well-being [1]. Efficient disease prevention hinges on possibility<br>59 are evaluating not only an individual's disease risk, but also the overall physiological vulnerability<br>50 in an early stage which i 59 evaluation only to the only stage which is often referred to as biological age. Originally the biological age of<br>51 individuals was estimated from a suite of physiological tests and biochemical clinical 60 in an early stage which is often referred to as biological age. Originally the biological age of<br>61 individuals was estimated from a suite of physiological tests and biochemical clinical<br>62 quantifications [2]. More rec 62 quantifications [2]. More recently explorations shifted towards comprehensive molecular<br>63 ('omics') datasets, providing global information on an individual's biological state. Currently 63 ('omics') datasets, providing global information on an individual's biological state. Currently<br>64 blood-based biomarkers to assess overall vulnerability in aging are constructed from 64 blood-based biomarkers to assess overall vulnerability in aging are constructed from<br>65 molecular markers and based on chronological age, disease onset and mortality [3]. Here we 65 molecular markers and based on chronological age, disease onset and mortality [3]. Here we<br>66 focus on data from metabolomics and proteomics platforms representing such novel focus on data from metabolomics and proteomics platforms representing such novel<br>67 molecular markers. 67 molecular markers.<br>68 focus on data from metabolomics and protein metabolomics and protein nuclear magnetic resonance (<sup>1</sup>HNMR) metabolomics enables a cost-effective and

12 molecular manusiers<br>68 Proton nuclear n<br>69 standardized assess Proton nuclear magnetic resonance (\*HNMR) metabolomics enables a cost-effective and<br>standardized assessment of a multitude of small circulating metabolites. Recent extensive<br>collaborative efforts, like BBMRI-NL [4], FINSK/ 69 Standardized assessment of a manufact of and a stating metabolites. Neverth antiting<br>70 collaborative efforts, like BBMRI-NL [4], FINSK/THL [5], COMETS [6], and the UK-Biobank [7],<br>71 resulted in large datasets generate The Teasulted in large datasets generated on the same Nightingale Health Pl <sup>1</sup>H-NMR<br>72 metabolomics platform. This platform has been largely explored as a source for generating resulted in large datasets generated on the same Nightingale Health Pl <sup>1</sup> 71 H-NMR T2 metabolomics platform. This platform has been largely explored as a source for generating<br>73 markers associated with a multitude of endpoints (e.g., type 2 diabetes [8], aging [4], risk<br>74 factors [9], and disease onset The manner assessment and a multitude of endpoint (e.g., type 2 diabetes [9], aging [4], risk<br>74 factors [9], and disease onset [10]). It gained particular attention, after training the<br>75 MetaboHealth score, using mortali 75 MetaboHealth score, using mortality as endpoint, in the largest study of its kind so far<br>76 (44,000 individuals and 5,500 incident deaths) [11]. This score stratifies mortality risk with a 75 MetaboHealth score, using mortality as endpoint, in the largest study of its kind so far<br>76 (44,000 individuals and 5,500 incident deaths) [11]. This score stratifies mortality risk with a<br>77 higher accuracy than conven higher accuracy than conventional clinical variables, with lower and higher values indicating 77 higher accuracy than conventional clinical variables, with lower and higher values indicating

178 low and higher 5 years risk for mortality, respectively. The MetaboHealth score, though<br>179 originally trained on mortality outcome, predicts multiple conditions related to overall<br>180 health decline associated with ag 78 health decline associated with ageing, such as frailty [12], cognitive decline [13], cancer, as<br>well as respiratory deficiencies [7]. Remarkably, the MetaboHealth score includes only 14 well as respiratory deficiencies [7]. Remarkably, the MetaboHealth score includes only 14<br>82 metabolic markers. These are involved in processes like glycolysis, fatty acid metabolism, 82 metabolic markers. These are involved in processes like glycolysis, fatty acid metabolism,<br>83 lipoproteins, and inflammation, e.g., GlycA. Although the MetaboHealth score offers an 82 metabolic markers. These are involved in processes like glycolysis, fatty acid metabolism,<br>83 lipoproteins, and inflammation, e.g., GlycA. Although the MetaboHealth score offers an<br>84 indication on physiological vulnera 84 indication on physiological vulnerability, especially for older individuals, it remains largely<br>85 illusive which pathophysiological mechanisms and corresponding blood factors are tracked 85 illusive which pathophysiological mechanisms and corresponding blood factors are tracked<br>86 by this mortality-trained risk score. 85 illusive which pathophysiological mechanisms and corresponding blood factors are tracked<br>86 by this mortality-trained risk score.<br>87 To address this question, we performed profiling of proteins and cytokines in

88 explore which molecular pathways co-vary with the MetaboHealth score. To this end, we explore which molecular pathways co-vary with the MetaboHealth score. To this end, we<br>89 tested for differences in plasma protein profiles of 50 out of 2200 Leiden Longevity Study tested for differences in plasma protein profiles of 50 out of 2200 Leiden Longevity Study<br>189 beta participants (mean age 156 y.o.) with the most extreme low and high values of tested for differences in plasma protein profiles of 50 out of 2200 Leiden Longevity Study<br>90 participants (mean age~ 56 y.o.) with the most extreme low and high values of<br>91 MetaboHealth; similarly 50 out of 2900 Rotterda 91 MetaboHealth; similarly 50 out of 2900 Rotterdam Study participants (mean age~ 67 y.o.),<br>92 and 50 monozygotic twins (MZTs) from the 25 most discordant MetaboHealth scoring pairs 92 and 50 monozygotic twins (MZTs) from the 25 most discordant MetaboHealth scoring pairs<br>93 out of 2,754 twins in the Netherlands Twin Register (mean age~ 36 y.o.). Considering that 93 out of 2,754 twins in the Netherlands Twin Register (mean age~ 36 y.o.). Considering that<br>94 MZTs have identical genomes, their within-pair associations are free of genetic confounding<br>95 [14]. Therefore, while the fir 93 out of 2,754 terms in the Netherlands Trum Register (mean age ~ 36 y.o.). Considering that<br>93 OUTS have identical genomes, their within-pair associations are free of genetic confounding<br>95 [14]. Therefore, while the fir 95 [14]. Therefore, while the first two cohorts offer an insight into population associations, the<br>96 MZTs differences inform to what extent MetaboHealth differences in the plasma proteins 96 MZTs differences inform to what extent MetaboHealth differences in the plasma proteins<br>97 and cytokines are unconfounded by shared genetics and environment. Overall, our<br>98 exploration leads to a better understanding of 97 and cytokines are unconfounded by shared genetics and environment. Overall, our<br>98 exploration leads to a better understanding of the predictive power of the MetaboHealth exploration leads to a better understanding of the predictive power of the MetaboHealth<br>99 score. exploration leads to a better understanding of the prediction of the prediction of the Metabo Health power of the

100 RESOLTS<br>101 **Descript**<br>102 We perfo 101 **Description of the dataset and study population**<br>102 We performed a nested case-control study design [15],<br>103 extreme MetaboHealth scores (MetaboHealth), 25 with hig 103 extreme MetaboHealth scores (MetaboHealth), 25 with high score and 25 with low score,<br>104 from the middle-aged cohort Leiden-Longevity-Study Partners-Offspring (LLS\_PAROFFS, 104 from the middle-aged cohort Leiden-Longevity-Study Partners-Offspring (LLS\_PAROFFS,<br>105 mean age~ 56 y.o.), and the Rotterdam Study (RS, mean age~ 67 y.o.) composed by older 105 mean age~ 56 y.o.), and the Rotterdam Study (RS, mean age~ 67 y.o.) composed by older<br>106 aged individuals (Figure 1B and Supplementary Table). To minimize potential confounding, mean age~ 56 y.o.), and the Rotterdam Study (RS, mean age~ 67 y.o.) composed by older<br>106 aged individuals (Figure 1B and Supplementary Table). To minimize potential confounding,<br>107 we ensured that the lower extreme sampl 107 we ensured that the lower extreme samples were age- and sex-matched, with at least one<br>108 high MetaboHealth case (Figure 1A-B). Given that a high MetaboHealth score corresponds to<br>109 higher mortality risk and poor he 108 high MetaboHealth case (Figure 1A-B). Given that a high MetaboHealth score corresponds to<br>109 higher mortality risk and poor health status, we categorized individuals with high scores as 109 higher mortality risk and poor health status, we categorized individuals with high scores as<br>110 "cases" and the remainder as "controls". These disparities in scores are manifested also in 110 "cases" and the remainder as "controls". These disparities in scores are manifested also in<br>111 phenotypic characteristics. We observed that the cases show significantly higher BMI in<br>112 LLS PAROFFS (cases: 26.53 vs 111 "phenotypic characteristics. We observed that the cases show significantly higher BMI in<br>112 LLS\_PAROFFS (cases: 26.53 vs controls: 24.37 kg/m<sup>2</sup>) and a higher incidence of 112 LLS\_PAROFFS (cases: 26.53 vs controls: 24.37 kg/m<sup>2</sup>) and a higher incidence of antihypertensive medication in RS (cases: 15 vs controls: 6 users) (Supplementary Table S1). LLS\_PAROFFS (cases: 26.53 vs controls: 24.37 kg/m<sup>2</sup>) and a higher incidence of<br>113 antihypertensive medication in RS (cases: 15 vs controls: 6 users) (Supplementary Table S1).<br>114 We observed in both study samples a lower 114 We observed in both study samples a lower level of lymphocyte percentage in the cases (RS:<br>115 cases= 25.58% vs controls= 35.24%, LLS\_PAROFFS: cases=24.72% vs controls= 31.52%) 2115 cases= 25.58% vs controls= 35.24%, LLS\_PAROFFS: cases=24.72% vs controls= 31.52%)<br>116 (Figure 1B). Interestingly, cases and controls shared similar phenotypic characteristics, 116 (Figure 1B). Interestingly, cases and controls shared similar phenotypic characteristics,<br>117 despite being derived from two independent cohorts, differing for the significantly higher 117 despite being derived from two independent cohorts, differing for the significantly higher<br>118 ages in RS (mean (age) <sub>RS</sub>= 74 y.o., mean (age) <sub>LLS-PAROFFs</sub>=59 y.o.), accompanied by a slightly 118 ages in RS (mean (age)  $_{RS}$  = 74 y.o., mean (age)  $_{LLS\text{-}PARCFFs}$ =59 y.o.), accompanied by a slightly<br>119 elevated BMI in RS (mean (BMI)  $_{RS}$ = 25.86, mean (BMI)  $_{LLS\text{-}PARCFFs}$ =24.86) and most 119 elevated BMI in RS (mean (BMI)  $_{RS}$ = 25.86, mean (BMI)  $_{LLS-PAROFFs}$ =24.86) and most<br>120 importantly the larger MetaboHealth contrast in RS (mean (MetaboHealth contrast)  $_{RS}$ = 2.4, 120 importantly the larger MetaboHealth contrast in RS (mean (MetaboHealth contrast)  $_{RS}$  = 2.4, mean (MetaboHealth contrast)  $_{LLS\text{-}PAROFFs}$ =1.7) (Figure 1B, Supplementary Table S1). 120 importantly the larger MetaboHealth contrast in RS (mean (MetaboHealth contrast)<br>121 mean (MetaboHealth contrast) <sub>LLS-PAROFFs</sub>= 1.7) (Figure 1B, Supplementary Table S1). 121 mean (Metabodhealth contrast) LLS-PAROFFS $=1.7$  (Figure 1B). Supplementary  $T$  and  $12$ ).

To investigate to what extent associations between proteins and MetaboHealth scores<br>123 reflect confounding by shared genetic or environmental factors, we investigated the contrast<br>124 in MetaboHealth discordant monozygoti 124 in MetaboHealth discordant monozygotic twins (MZTs) from the Netherland Twin Register<br>125 dataset (NTR, mean age~ 36 y.o.) (Figure 1A). 25 Twin pairs with the highest discordance 125 dataset (NTR, mean age~ 36 y.o.) (Figure 1A). 25 Twin pairs with the highest discordance<br>126 with respect to their MetaboHealth score were selected from 2,754 NTR participants. In 126 vith respect to their MetaboHealth score were selected from 2,754 NTR participants. In<br>127 accordance with the observations in RS and LLS, the individuals with higher MetaboHealth 127 accordance with the observations in RS and LLS, the individuals with higher MetaboHealth<br>128 exhibited a significant reduction in the lymphocyte percentage (cases: 31.37% vs controls: 2128 exhibited a significant reduction in the lymphocyte percentage (cases: 31.37% vs controls:<br>129 38.64% ). Nonetheless, the NTR shows some relevant characteristics when compared to the 129 38.64%). Nonetheless, the NTR shows some relevant characteristics when compared to the<br>130 other two studies. It represents the youngest population, with a 20-year gap 129 38.64% ). Nonetheless, the NTR shows some relevant characteristics when compared to the<br>130 other two studies. It represents the youngest population, with a 20-year gap<br>131 (mean(age)<sub>NTR</sub>=36 y.o.), it has a larger pr 131 (mean(age)<sub>NTR</sub>=36 y.o.), it has a larger presence of females (36 out of 50), for which the<br>132 MetaboHealth previously showed a reduced accuracy [16]. These differences, and the 132 MetaboHealth previously showed a reduced accuracy [16]. These differences, and the<br>133 selection criterium based on the twin discordancy, possibly lead to a diminished contrast in 132 MetaboHealth previously showed a reduced accuracy [16]. These differences, and the<br>selection criterium based on the twin discordancy, possibly lead to a diminished contrast in<br>MetaboHealth values (mean (MetaboHealth c MetaboHealth values (mean (MetaboHealth contrast) NTR= 1.18) (Figure B).

### 136 Luminex cytokine assays: higher levels of GDF15, IL6, and MIG in the

### 135

136 Luminex cytokine assays: higher levels of GDF15, IL6, and MIG in the<br>137 **participants with a high MetaboHealth score**<br>138 To explore the inflammatory state variation between the MetaboHealth extremes, we 137 **participants with a high Metabonealth score**<br>138 To explore the inflammatory state variation betw<br>139 quantified 15 cytokines on the Luminex platform (M To explore the inflammatory state variation between the MetaboHealth extremes, we<br>
quantified 15 cytokines on the Luminex platform (Materials and Methods for a complete<br>
explanation). Six out of the 15 (40%) cytokines (IL2 below the detection threshold in most samples, likely implying that the participants were 141 below the detection threshold in most samples, likely implying that the participants were<br>142 relatively healthy at the time of sampling (Figure 2A and S1). Indeed, on average, more 142 relatively healthy at the time of sampling (Figure 2A and S1). Indeed, on average, more<br>143 cytokines go undetected in low MetaboHealth participants across the LLS and RS cohorts, 142 relatively healthy at the time of sampling (Figure 2A). This 22). Indeed, on average, more<br>143 relatively provided in low MetaboHealth participants across the LLS and RS cohorts,<br>144 hinting at their overall lower infl 144 hinting at their overall lower inflammation rate (Figure 2A). This discrepancy in detectability<br>144 hinting at their overall lower inflammation rate (Figure 2A). This discrepancy in detectability 144 hinting at their overall lower inflammation rate (Figure 2A). This discrepancy in detectability

attains statistical significance in the case of IL6 (p-value=0.001) (Figure 2A). The same<br>146 distinctive patterns between the high and low MetaboHealth score cohorts can be observed<br>147 when considering the two cohorts se

Our subsequent analyses focused on nine cytokines (MIP1a, IL6, RANTES, MIG, MCP1, 148 Our subsequent analyses focused on nine cytokines (MI<br>149 Eotaxin, Psel, GDF15, and BDNF) exhibiting the fewest deter 148 Eotaxin, Psel, GDF15, and BDNF) exhibiting the fewest detectability issues (Figure 2A). We<br>150 assessed differential expression of these cytokines between MetaboHealth cases and 149 Eotaxin, Psel, GDF15, and BDNF) exhibiting the fewest detectability issues (Figure 2A). We<br>150 assessed differential expression of these cytokines between MetaboHealth cases and<br>151 controls, based on a univariate line 152 Interestingly, significantly higher levels of GDF15 (estimate~1.08, fdr=3.9 x 10<sup>-8</sup>), IL6 152 Interestingly, significantly higher levels of GDF15 (estimate~1.08, fdr=3.9 x  $10^{-8}$ ), IL6<br>153 (estimate~1.05, fdr=1.8 x  $10^{-4}$ ), and MIG (estimate~0.95, fdr=1.23 x  $10^{-3}$ ) were observed in Interestingly, significantly higher levels of GDF15 (estimate~1.08, fdr=3.9 x 10°), IL6<br>153 (estimate~1.05, fdr=1.8 x 10<sup>-4</sup>), and MIG (estimate~0.95, fdr=1.23 x 10<sup>-3</sup>) were observed in<br>154 the high MetaboHealth group wh (estimate~1.05, fdr=1.8 x 10<sup>-</sup>), and MIG (estimate~0.95, fdr=1.23 x 10<sup>-</sup>) were observed in<br>154 the high MetaboHealth group when considering LLS\_PAROFFS and RS together (Figure 2B).<br>155 Adjusting for medication usag 155 Adjusting for medication usage (blood pressure lowering and statins) and cell count<br>156 (particularly lymphocyte %) influenced mostly the association with GDF15 (Figure S2A-D), 156 (particularly lymphocyte %) influenced mostly the association with GDF15 (Figure S2A-D),<br>157 yet it remained significant. Finally, reproducing the univariate associations separately for the 157 (particularly left interest of the univariate associations separately for the two cohorts shows similar patterns (Figure S2F), underpinning their robustness. 158 two cohorts shows similar patterns (Figure S2F), underpinning their robustness.<br>159 To further investigate the origin of the observed signal, we looked into the correlation

159 To further investigate the origin of the observed signal, we looked into the<br>160 structure between cytokines. The generally modest intercorrelation showed 159 To further investigate the origin of the observed signal, we looked into the correlation<br>160 structure between cytokines. The generally modest intercorrelation showed a profile of<br>161 mostly independent features (Figur 161 mostly independent features (Figure S2E). Next, we estimated the correlation of the cytokines with the metabolomics components of the MetaboHealth score (Figure 2C). While 162 cytokines with the metabolomics components of the MetaboHealth score (Figure 2C). While<br>163 MetaboHealth exhibits the highest correlations, we observed several noteworthy relations, 163 MetaboHealth exhibits the highest correlations, we observed several noteworthy relations,<br>164 i.e. strong positive correlations between Glycoprotein Acetyls (GlycA), a metabolomics<br>165 inflammation marker, and GDF15 ( 164 i.e. strong positive correlations between Glycoprotein Acetyls (GlycA), a metabolomics<br>165 inflammation marker, and GDF15 (r=0.33, p=8.2 x 10<sup>-4</sup>), MIG (r=0.27, p= 7.5x 10<sup>-3</sup>), and IL6 165 inflammation marker, and GDF15 (r=0.33, p=8.2 x 10<sup>-4</sup>), MIG (r=0.27, p= 7.5x 10<sup>-3</sup>), and IL6 (r=0.33, p= 1.9x 10<sup>-3</sup>). Intriguingly, GDF15 displayed also elevated positive correlations with inflammation marker, and GDF15 (r=0.33, p=8.2 x 10<sup>-</sup>), MIG (r=0.27, p= 7.5x 10<sup>-</sup>), and IL6<br>166 (r=0.33, p= 1.9x 10<sup>-3</sup>). Intriguingly, GDF15 displayed also elevated positive correlations with<br>167 glucose (r=0.35, p= 4.2 166 (r=0.33, p= 1.9x 10<sup>-3</sup>). Intriguingly, GDF15 displayed also elevated positive correlations with<br>167 glucose (r=0.35, p= 4.2x 10<sup>-4</sup>), phenylalanine (r=0.35, p= 3.0 x 10<sup>-3</sup>), and isoleucine (r=0.23,<br>168 p= 2.1x 10<sup>-2</sup> glucose (r=0.35, p= 4.2x  $10^4$ ), phenylalanine (r=0.35, p= 3.0 x 10<sup>-3</sup>), and isoleucine (r=0.23,<br>168 p= 2.1x 10<sup>-2</sup>). In contrast, we uncovered prominent negative correlations between GDF15 p= 2.1x 10 °). In contrast, we uncovered prominent negative correlations between GDF15 ).<br>In contrast, we uncovered prominent negative correlations between GDF15 .<br>In contrast, we uncovered prominent negative correlations

and S-HDL-L (total lipids in small HDL) (r=-0.38, p= 9.9x 10<sup>-</sup>), and between MIG and Histidine<br>170 (r=-0.33, p= 8.7x 10<sup>-4</sup>).<br>171

## 170 (r=-0.33, p= 8.7x 10 <sup>-</sup>).<br>171<br>172 **Plasma proteon** 171<br>172<br>173 **Plasma proteomics associated with extremes in MetaboHealth are**<br>173 **enriched for inflammatory response and cholesterol transport pathways**<br>174 In the two population-based samples (50 cases and 50 controls), we investigat

173 Enriched for inflammatory response and cholesteror transport pathways<br>174 In the two population-based samples (50 cases and 50 controls), we investigated p<br>175 proteome profiles by a DIA-based quantitative plasma prote 175 proteome profiles by a DIA-based quantitative plasma proteomics pipeline [17,18]<br>176 (Materials and Methods). 261 out of 337 measured plasma proteins (77%) passed the proteome profiles by a DIA-based quantitative plasma proteomics pipeline [17,18]<br>176 (Materials and Methods). 261 out of 337 measured plasma proteins (77%) passed the<br>177 detection limit and quality control criteria (detai 177 detection limit and quality control criteria (detailed in Materials and Methods). We<br>178 identified 106 (68 negative and 38 positive) significant univariate linear associations with the 178 identified 106 (68 negative and 38 positive) significant univariate linear associations with the<br>179 MetaboHealth extremes, adjusted for age, sex, and BMI (Figure 3A). APOA1 (estimate~-1.46, 179 MetaboHealth extremes, adjusted for age, sex, and BMI (Figure 3A). APOA1 (estimate~-1.46,<br>180 fdr=2.44 x 10<sup>-13</sup>), APOA2 (estimate~ -1.4, fdr= 4.02x 10<sup>-12</sup>), TETN (estimate~-1.31, fdr= 6.42x 180 fdr=2.44 x 10<sup>-13</sup>), APOA2 (estimate~ -1.4, fdr= 4.02x 10<sup>-12</sup>), TETN (estimate~-1.31, fdr= 6.42x<br>181 10<sup>-11</sup>), GELS (estimate~-1.26, fdr= 3.18 x 10<sup>-10</sup>), and APOA4 (estimate~ -0.93, fdr= 5.71 x 10<sup>-6</sup>) fdr=2.44 x 10<sup>-13</sup>), APOA2 (estimate~ -1.4, fdr= 4.02x 10<sup>-12</sup>), TETN (estimate~-1.31, fdr= 6.42x<br>181 10<sup>-11</sup>), GELS (estimate~-1.26, fdr= 3.18 x 10<sup>-10</sup>), and APOA4 (estimate~ -0.93, fdr= 5.71 x 10<sup>-6</sup>)<br>182 were the stro 181 10<sup>-1</sup>), GELS (estimate~-1.26, fdr= 3.18 x 10<sup>-0</sup>), and APOA4 (estimate~ -0.93, fdr= 5.71 x 10<sup>-0</sup>)<br>182 were the strongest negative associated proteins. In contrast, the positive acute phase<br>183 proteins CRP (estimate 183 proteins CRP (estimate~1.51, fdr= 5.81x 10<sup>-14</sup>), LBP (estimate~1.35, p= 1.88 x 10<sup>-11</sup>), HPT<br>184 (estimate~1.14, p= 1.74 x 10<sup>-8</sup>), were the strongest positive associated proteins. In line with proteins CRP (estimate~1.51, fdr= 5.81x 10<sup>-14</sup>), LBP (estimate~1.35, p= 1.88 x 10<sup>-11</sup>), HPT<br>184 (estimate~1.14, p= 1.74 x 10<sup>-8</sup>), were the strongest positive associated proteins. In line with<br>185 what was observed for (estimate~1.14, p= 1.74 x 10<sup>-o</sup>), were the strongest positive associated proteins. In line with<br>185 what was observed for the cytokines, the additional correction for medication usage had<br>186 minimal impact on the 186 minimal impact on the univariate associations, while lymphocyte percentage exhibited a<br>187 slightly more pronounced influence (Figure S4B-D). Furthermore, a meta-analyses to 187 slightly more pronounced influence (Figure S4B-D). Furthermore, a meta-analyses to<br>188 evaluate the associations separately in the two cohorts revealed alike results, with generally 188 evaluate the associations separately in the two cohorts revealed alike results, with generally<br>189 stronger signals in the RS (Figure S5 A-B).

190 Subsequently, we explored the statistical interrelations and biological functionalities of<br>191 the 106 plasma proteome features that exhibited significant association to the 189 Subsequently, we explored the statist<br>191 the 106 plasma proteome features 191 the 106 plasma proteome features that exhibited significant association to the 191 the 106 plasma proteome features that exhibited significant association to the

MetaboHealth extremes. These proteins showed high correlations with GlycA, comparable<br>193 to those observed for the MetaboHealth score (Figure S6A). Moreover, within the group of<br>194 positively associated proteins, smaller positively associated proteins, smaller clusters of highly correlated proteins are found (Figure 195 S6B). To gain some biological interpretation, we employed KEGG and Gene Ontology to<br>196 perform functional enrichments separately for the positively and negatively associated 196 perform functional enrichments separately for the positively and negatively associated<br>197 proteins (Figure S7). As expected, considering that lower MetaboHealth values are related to 196 perform functional enrichments separately for the positively and negatively associated<br>197 proteins (Figure S7). As expected, considering that lower MetaboHealth values are related to<br>198 healthier metabolic profiles, 198 healthier metabolic profiles, the negatively associated proteins demonstrated high<br>199 enrichments for processes relating to "cholesterol transport", "cholesterol metabolism", and 199 enrichments for processes relating to "cholesterol transport", "cholesterol metabolism", and<br>200 "high-density lipoproteins particle remodeling" (Figure 3B, S7C and S7E). Conversely, the enrichments for processes relating to "cholesterol transport", "cholesterol metabolism", and<br>
200 Thigh-density lipoproteins particle remodeling" (Figure 3B, S7C and S7E). Conversely, the<br>
201 positively associated protein positively associated proteins were more enriched with "Inflammatory response",<br>complement and coagulation cascades", and intriguingly, "Coronavirus disease" (Figure 3B, 202 "complement and coagulation cascades", and intriguingly, "Coronavirus disease" (Figure 3B,<br>203 S7D and S7F). The latter can be interpreted as a validation, as a subset of about a dozen of 203 S7D and S7F). The latter can be interpreted as a validation, as a subset of about a dozen of<br>204 inflammation related features, measured with the same proteomics platform, were 203 S7D and S7F). The latter can be interpreted as a validation, as a subset of about a dozen of<br>204 inflammation related features, measured with the same proteomics platform, were<br>205 previously linked to COVID19 outcome 205 previously linked to COVID19 outcome [18]. Four of these twelve COVID related plasma<br>206 proteins exhibited consistently statistically significant differences between the extremes of 206 proteins exhibited consistently statistically significant differences between the extremes of<br>207 the MetaboHealth score in RS and LLS-PAROFFs (Figure S8). 207 blues exhibited consistently statistically significant differences between the extremes of the MetaboHealth score in RS and LLS-PAROFFs (Figure S8).

## 207 the MetaboHealth score in RS and LLS-PAROFFs (Figure S8).<br>208<br>209 **Genetic influences on the MetaboHealth score and analysis of plasma** 209<br>210 **210 proteins in extreme discordant MZ twins**<br>211 The NTR has collected 726 complete monozygotic twin (MZTs) and 450 dizygotic twins

210 **proteins in extreme discordant MZ twins**<br>211 The NTR has collected 726 complete monozy<br>212 (DZTs) twin pairs with metabolomics data. We es 212 (DZTs) twin pairs with metabolomics data. We estimated the resemblance in MetaboHealth<br>213 core as a function of zygosity in these pairs. The correlation between MZ twin pairs was 213 score as a function of zygosity in these pairs. The correlation between MZ twin pairs was<br>214 estimated as  $r = 0.432$  (95% CI = 0.370-0.489), and the correlation in DZTs was  $r = 0.230$  (95% 214 schemated as  $r = 0.432$  (95% CI = 0.370-0.489), and the correlation in DZTs was  $r = 0.230$  (95% 214 estimated as r= 0.432 (95% CI = 0.370-0.489), and the correlation in DZTs was r = 0.230 (95% CI = 0.141-0.316), indicating the MetaboHealth score as a heritable trait (h<sup>2</sup> = 0.4) (detailed<br>216 = information in Materials and Methods) [19].<br>217 = To exclude potential confounding from genetic factors within our asso

217 To exclude potential confounding from<br>218 conducted a monozygotic twin (MZTs) discor 218 conducted a monozygotic twin (MZTs) discordant twin pairs design; i.e., high and low scoring<br>219 MetaboHealth genetically matched twins. In this design an observed effect is not 219 MetaboHealth genetically matched twins. In this design an observed effect is not<br>220 confounded by genetic factors. Therefore, from the total population of MZ twin pairs, we 219 MetaboHealth genetically matched twins. In this design an observed effect is not<br>220 confounded by genetic factors. Therefore, from the total population of MZ twin pairs, we<br>221 selected a subsample of 25 most discorda 221 selected a subsample of 25 most discordant MZ twin pairs to further explore associations of<br>222 the score to the Luminex based cytokines and Mass Spectrometry-base proteomics profiling. 222 the score to the Luminex based cytokines and Mass Spectrometry-base proteomics profiling.<br>223 Concordantly with the previous sections, the protein markers with the strongest associations the score to the Luminex based cytokines and Mass Spectrometry-base proteomics profiling.<br>223 Concordantly with the previous sections, the protein markers with the strongest associations<br>224 with the MetaboHealth score als 225 NTR dataset, albeit less than in the LLS and RS studies (Figure 4, and S11). To take advantage 225 NTR dataset, albeit less than in the LLS and RS studies (Figure 4, and S11). To take advantage<br>226 of the genetic similarity of the MZ individuals, we tested for associated proteins using a 226 of the genetic similarity of the MZ individuals, we tested for associated proteins using a<br>227 Ilinear mixed model (Methods). The cytokines did not show significant differences in this 227 Inear mixed model (Methods). The cytokines did not show significant differences in this<br>228 vithin-pairs design, although we observed similar trends to the results in LLS and RS, with 228 within-pairs design, although we observed similar trends to the results in LLS and RS, with<br>229 elevated cytokines in twins with high MetaboHealth scores (Figure S10A). The proteomics within-pairs design, although we observed similar trends to the results in LLS and RS, with<br>229 elevated cytokines in twins with high MetaboHealth scores (Figure S10A). The proteomics<br>230 analyses revealed a robust signal, 220 analyses revealed a robust signal, identifying a total of 86 significant associations (Figure S10B). Notably, CRP (estimate~1.19, fdr= 3.75x  $10^5$ ) and LBP (estimate~1.15, fdr= 8.39x  $10^5$ ) 231 S10B). Notably, CRP (estimate~1.19, fdr= 3.75x 10<sup>-5</sup>) and LBP (estimate~1.15, fdr= 8.39x 10<sup>-5</sup>) once again emerged as the most prominently positively associated proteins, while TETN S10B). Notably, CRP (estimate~1.19, fdr= 3.75x 10<sup>-3</sup>) and LBP (estimate~1.15, fdr= 8.39x 10<sup>-3</sup>)<br>232 once again emerged as the most prominently positively associated proteins, while TETN<br>233 (estimate~-1.2, fdr= 9.69x 10 233 (estimate~-1.2, fdr= 9.69x 10<sup>5</sup>) and GELS (estimate~-1.11, fdr= 4.43x 10<sup>5</sup>) were confirmed as the most negatively associated proteins. Moreover, the lower associations of APOA1 and 235 APOA2 in the twins' profiles d (estimate~-1.2, fdr= 9.69x 10<sup>-</sup>) and GELS (estimate~-1.11, fdr= 4.43x 10<sup>-</sup>) were confirmed as<br>234 in the most negatively associated proteins. Moreover, the lower associations of APOA1 and<br>235 APOA2 in the twins' profile 235 APOA2 in the twins' profiles differs from the contrast in LLS and RS and suggests the<br>236 presence of genetic confounding in the associations between these proteins and the presence of genetic confounding in the associations between these proteins and the<br>237 MetaboHealth scores. A correction for health factors had similar results as for the other two<br>238 cohorts, with lymphocyte percentage h MetaboHealth scores. A correction for health factors had similar results as for the other two cohorts, with lymphocyte percentage having the highest effect (Figure S10C-D). 238 cohorts, with lymphocyte percentage having the highest effect (Figure S10C-D).



**DISCUSSION**<br>248 The MetaboHealth score, along with other <sup>1</sup>H-NMR metabolomics-based markers,<br>249 displays risk stratification across a spectrum of health and disease outcomes relevant in The MetaboHealth score, along with other "H-NMR metabolomics-based markers,<br>249 Heisplays risk stratification across a spectrum of health and disease outcomes relevant in<br>250 Hageing research. The score, though based on mo 250 ageing research. The score, though based on mortality, has shown to be an indicator of overall health status in middle and older aged individuals. Within this context, our study set ageing research. The score, though based on mortality, has shown to be an indicator of<br>251 overall health status in middle and older aged individuals. Within this context, our study set<br>252 out to quantify comprehensive pl 252 out to quantify comprehensive plasma proteome profiles in 150 samples at the extreme<br>253 ends of the MetaboHealth distributions from three large Dutch cohorts (Leiden Longevity<br>254 Study, Rotterdam Study, and the Nethe 252 out to quantify comprehensive plasma processive profiles in 250 samples at the extreme<br>253 out of the MetaboHealth distributions from three large Dutch cohorts (Leiden Longevity<br>254 Study, Rotterdam Study, and the Neth 254 Study, Rotterdam Study, and the Netherlands Twin Register, spanning a total dataset of<br>255 - 7,854 individuals). Our findings revealed significant differential expression among 106 255 7,854 individuals). Our findings revealed significant differential expression among 106<br>256 plasma proteins and 3 cytokines markers, consistently observed in the RS and LLS-PAROFFs,<br>257 between the highest (cases) and 256 plasma proteins and 3 cytokines markers, consistently observed in the RS and LLS-PAROFFs,<br>257 between the highest (cases) and the lowest (controls) MetaboHealth scores, respectively 257 between the highest (cases) and the lowest (controls) MetaboHealth scores, respectively<br>258 indicating elevated and reduced mortality risk. These associations were for the majority not 258 indicating elevated and reduced mortality risk. These associations were for the majority not<br>259 confounded by age, sex, BMI, and medication usage. A part of the proteins associated with<br>260 the MetaboHealth contrast c 259 confounded by age, sex, BMI, and medication usage. A part of the proteins associated with<br>260 the MetaboHealth contrast could be explained by genetic confounding as demonstrated by 269 confounded by and measure usage. A part of the proteins and the MetaboHealth contrast could be explained by genetic confounding as demonstrated by investigating discordant monozygotic twins.

261 investigating discordant monozygotic twins.<br>262 The majority of the significant associations (68 out of 106) resulted in negatively<br>263 associated proteins with the case/control contrast, indicating higher protein leve 263 associated proteins with the case/control contrast, indicating higher protein levels in<br>264 samples of healthier subjects (reflected by lower MetaboHealth scores). Functional 263 associated proteins with the case/control contrast, indicating higher protein levels in<br>264 samples of healthier subjects (reflected by lower MetaboHealth scores). Functional 264 samples of healthier subjects (reflected by lower MetaboHealth scores). Functional<br>265 enrichment of these proteins revealed associations with healthy metabolism, particularly in 265 enrichment of these proteins revealed associations with healthy metabolism, particularly in<br>266 high-density lipoprotein (HDL) remodeling and cholesterol transport pathways. Notably, the<br>267 most prominent associated p 266 high-density lipoprotein (HDL) remodeling and cholesterol transport pathways. Notably, the<br>267 most prominent associated proteins were APOA1 and APOA2, crucial components of HDL 267 most prominent associated proteins were APOA1 and APOA2, crucial components of HDL<br>268 and widely recognized as protective markers for cardiovascular disease [20,21]. It should be 268 and widely recognized as protective markers for cardiovascular disease [20,21]. It should be<br>269 noted that likely a substantial genetic component drives APOA1 and 2b, since these proteins noted that likely a substantial genetic component drives APOA1 and 2b, since these proteins 269 noted that likely a substantial genetic component drives APOA1 and 2b, since these proteins

270 were significantly less discordant in the MetaboHealth discordant MZ twins. This aspect<br>271 must be further investigated since a genetic component may be overestimated given that<br>272 the MetaboHealth contrast in the (overa the MetaboHealth contrast in the (overall younger) MZ twins was rather small in comparison 273 to the cohort studies. Interestingly, the associations with tetranectin (TETN) and gelsolin<br>274 (GELS) emerged as consistently stable also in NTR. Both these proteins are under 274 (GELS) emerged as consistently stable also in NTR. Both these proteins are under<br>275 consideration as potential protective markers for various diseases, such as cancer, 274 (GELS) emerged as consistently stable also in NTR. Both these proteins are under<br>275 consideration as potential protective markers for various diseases, such as cancer,<br>276 cardiovascular-disease-and-neurodegeneration

277 Consideration as exert mate variation as the verest matter in the also clated various VII various 277 Considers Musulm as cancers we also discussed as 277 Considered as 277 Considered as 277 Consideration as cancer we 277 CME also observed several relevant proteins that we<br>278 CME With the MetaboHealth contrast. The cytokine mar 278 with the MetaboHealth contrast. The cytokine markers selected for this study were<br>279 previously reported to strongly associate with frailty, ageing, age-related disease, and 278 with the MetaboHealth contrast. The cytokine markers selected for this study were<br>279 previously reported to strongly associate with frailty, ageing, age-related disease, and<br>280 mortality [25–28]. While a large part o 280 mortality [25–28]. While a large part of the cytokines could not be efficiently detected,<br>281 which can be attributed to the absence of ongoing infections and acute inflammation at 281 which can be attributed to the absence of ongoing infections and acute inflammation at<br>282 blood sampling, GDF15, IL6, and MIG (gene name=CXCL9), showed significant associations which can be attributed to the absence of ongoing infections and acute inflammation at<br>282 blood sampling, GDF15, IL6, and MIG (gene name=CXCL9), showed significant associations<br>283 with higher MetaboHealth levels. These c 282 blood sampling, GDF15, 129, and MIF (gene name CRCL9), showed significant associations<br>284 blood significant associations and MIG (gene name significant roles in the body and are<br>284 blood in cell signalling immune res 284 involved in cell signalling immune response, and inflammation. All three markers are<br>285 potential biomarkers for aging-related physiological decline and frailty, where IL6 marks 285 potential biomarkers for aging-related physiological decline and frailty, where IL6 marks<br>286 chronic inflammation ("inflammaging") [29–32]; GDF-15 is a senescence associated<br>287 secretory protein (SASP) and a marker o 286 chronic inflammation ("inflammaging") [29–32]; GDF-15 is a senescence associated secretory protein (SASP) and a marker of physiological stress response and mitochondrial 287 secretory protein (SASP) and a marker of physiological stress response and mitochondrial<br>288 dysfunction and MIG is the most prominent marker emerging from the inflammatory aging 288 dysfunction and MIG is the most prominent marker emerging from the inflammatory aging<br>289 clock (see materials and Methods) [25]. In this regard, also the 38 significantly positively<br>290 associated plasma proteins are 289 clock (see materials and Methods) [25]. In this regard, also the 38 significantly positively<br>290 associated plasma proteins are predominantly involved in inflammatory response, 280 associated plasma proteins are predominantly involved in inflammatory response,<br>291 complement and coagulation cascades, and COVID19. CRP, HPT and LBP, well-known 291 complement and coagulation cascades, and COVID19. CRP, HPT and LBP, well-known<br>292 biomarkers assessing the degree of inflammation and immune response, displayed the<br>293 highest significant associations. biomarkers assessing the degree of inflammation and immune response, displayed the 293 bighest significant associations. 293 highest significant associations.

295 to COVID19 infection pathways, regardless of their infection status, considering that blood<br>296 Samples used in this study were drawn up to 15 years before the COVID19 pandemic. 296 samples used in this study were drawn up to 15 years before the COVID19 pandemic.<br>297 Noteworthy is that both the proteomics panel and the metabolomics assay composing the 296 samples used in this study were drawn up to 15 years before the COVID19 pandemic.<br>297 Noteworthy is that both the proteomics panel and the metabolomics assay composing the<br>298 MetaboHealth score in plasma, were previou 298 MetaboHealth score in plasma, were previously observed to be able to stratify severe cases<br>299 of COVID19 [7,18]. Several plasma proteins previously related to COVID19 were significantly 299 of COVID19 [7,18]. Several plasma proteins previously related to COVID19 were significantly<br>300 different also in our MetaboHealth contrast in at least one study. These include the inter- $\alpha$ -200 different also in our MetaboHealth contrast in at least one study. These include the inter- $\alpha$ -<br>201 trypsin inhibitors family, ITIH1, ITIH2, respectively positively and negatively associated, and 301 trypsin inhibitors family, ITIH1, ITIH2, respectively positively and negatively associated, and<br>302 HRG, LCAT, which are both positively associated to MetaboHealth (Figure S8A). 301 trypsin inhibitors family, ITIH1, ITIH2, respectively positively and negatively associated, and<br>302 HRG, LCAT, which are both positively associated to MetaboHealth (Figure S8A).<br>303 Significant disparities in phenotypi

303 Significant disparities in phenotypic characteristics were observed betwe<br>304 controls across the cohorts, mostly in the higher levels of BMI and antihyper 304 controls across the cohorts, mostly in the higher levels of BMI and antihypertensive usage<br>305 Fespectively in LLS and RS. However, of great importance from a health perspective seemed 304 controls across the cohorts, mostly in the higher levels of BMI and antihypertensive usage<br>305 respectively in LLS and RS. However, of great importance from a health perspective seemed<br>306 to be the significant reducti 307 studies, although still within healthy ranges (20-40%). Concordantly, the systematic<br>308 adjustment for health and risk-factors-(age,-sex,-BMI,-lymphocyte and-monocyte %,-lipid-and-307 studies, although still within healthy ranges (20-40%). Concordantly, the systematic<br>308 adjustment for health and risk factors (age, sex, BMI, lymphocyte and monocyte %, lipid and 308 adjustment for health and risk factors (age, sex, BMI, lymphocyte and monocyte %, lipid and<br>309 antihypertensive medication usage) resulted in an attenuation of the signal to 89/106 309 antihypertensive medication usage) resulted in an attenuation of the signal to 89/106<br>310 plasma proteins and 3/3 cytokines, with the strongest effect being related to the lymphocyte<br>311 percentage (Figure S4A-D). Inte percentage (Figure S4A-D). Interestingly, a decrease in lymphocyte counts and increased 311 percentage (Figure S4A-D). Interestingly, a decrease in lymphocyte counts and increased<br>312 cytokine in blood is frequently related to a decline in immune system functions that 312 cytokine in blood is frequently related to a decline in immune system functions that<br>313 accompanies physiological aging [33–36]. Considering the age matching within our study 313 accompanies physiological aging [33–36]. Considering the age matching within our study<br>314 inclusion criteria, these observations suggest that the MetaboHealth score successfully 314 inclusion criteria, these observations suggest that the MetaboHealth score successfully<br>315 identified individuals with accelerated biological aging given their chronological age. While 315 identified individuals with accelerated biological aging given their chronological age. While<br>316 the correction for cell counts partially accounted for the observed signal, it was evident that 316 the correction for cell counts partially accounted for the observed signal, it was evident that<br>317 hematopoietic variation alone does not explain the observed MetaboHealth contrast in 316 the correction for cell counts partially accounted for the observed signal, it was evident that<br>317 hematopoietic variation alone does not explain the observed MetaboHealth contrast in 317 hematopoietic variation alone does not explain the observed MetaboHealth contrast in

319 MetaboHealth and cell count percentages can jointly be more informative in indicating<br>320 health status of older individuals and secondly what the relation with these parameters and 320 health status of older individuals and secondly what the relation with these parameters and<br>321 inflammation is... 320 health status of older individuals and secondly what the relation with these parameters and<br>321 inflammation is. .<br>322 Conducting heritability analyses within the Netherlands Twin Register allowed us to

322 Conducting he<br>323 establish that the 323 establish that the MetaboHealth score has a heritability of ~40%, at least in relatively<br>324 younger ages (mean age~36 y.o.). Therefore, the Monozygotic twin subset provided an ideal ounger ages (mean age~36 y.o.). Therefore, the Monozygotic twin subset provided an ideal<br>325 - setting to further examine the proteomics associations within a genetically identical 325 setting to further examine the proteomics associations within a genetically identical<br>326 population. Cytokines did not exhibit significant associations to the MetaboHealth contrast 325 setting to further examine the proteomics associations within a genetically identical<br>326 population. Cytokines did not exhibit significant associations to the MetaboHealth contrast<br>327 in the twins, while we found up 327 in the twins, while we found up to 46 negative and 22 positive plasma protein relations. An<br>328 attenuation of the number and strength of associations indicate that part of the signal may 328 attenuation of the number and strength of associations indicate that part of the signal may<br>329 be explained by genetic factors that influence both the MetaboHealth score and quantified 329 be explained by genetic factors that influence both the MetaboHealth score and quantified<br>330 proteins. Further research should be focused on investigating the genetic confounding in be explained by genetic factors that influence both the MetaboHealth score and quantified<br>330 proteins. Further research should be focused on investigating the genetic confounding in<br>331 twins of higher ages where MetaboHe 1331 twins of higher ages where MetaboHealth is even more informative. Pertaining to this, it is<br>332 essential to highlight that NTR participants had lower MetaboHealth contrasts as they are 332 essential to highlight that NTR participants had lower MetaboHealth contrasts as they are<br>333 younger of age as well as due to the inclusion restriction. Supposedly for the same reasons, younger of age as well as due to the inclusion restriction. Supposedly for the same reasons,<br>334 diminished associations are noted also when comparing the results in LLS-PAROFFs to RS<br>335 (Figures S5). Possibly as individu (Figures S5). Possibly as individuals age, environmental and/or stochastic factors gain 335 (Figures S5). Possibly as individuals age, environmental and/or stochastic factors gain<br>336 greater importance on the features included in the MetaboHealth score over genetic greater importance on the features included in the MetaboHealth score over genetic<br>337 influences. greater importance on the features included in the MetaboHealth score over genetic<br>337 influences.<br>338 One aspect of our study that can be regarded as a limitation is the generally healthy state<br>339 of the participants in

338 One asp 339 of the participants in the cohorts from which we derived our participants. Consequently, the<br>340 observed contrast between the extreme levels of MetaboHealth was relatively modest. This<br>341 is underscored by the fact t 340 observed contrast between the extreme levels of MetaboHealth was relatively modest. This<br>341 is underscored by the fact that several of the examined cytokines consistently fell under the 341 is underscored by the fact that several of the examined cytokines consistently fell under the<br>341 is underscored by the fact that several of the examined cytokines consistently fell under the 341 is underscored by the fact that several of the examined cytokines consistently fell under the

detection limit, a circumstance that can be largely attributed to the overall absence of<br>343 ongoing infections at blood sampling which proteins may be useful biomarkers in clinical<br>344 studies. We envision MetaboHealth as studies. We envision MetaboHealth as a score potentially giving an indication of vulnerability<br>in individuals from the population in a modifiable health phase, way before onset and 345 in individuals from the population in a modifiable health phase, way before onset and<br>346 diagnosis of specific diseases. The score indeed effectively identified relevant biological 346 diagnosis of specific diseases. The score indeed effectively identified relevant biological<br>347 differences within all three populations indicating that early changes in multiple metabolic diagnosis of specific diseases. The score indeed effectively identified relevant biological<br>347 differences within all three populations indicating that early changes in multiple metabolic<br>348 and proteome parameters known 348 and proteome parameters known to reflect decline in health is represented by the score.<br>349 Given the limited number of individuals in our selection we do not have the statistical power 349 Given the limited number of individuals in our selection we do not have the statistical power<br>350 to explore concrete endpoints such as mortality or frailty in this study. We consider this 349 Given the limited number of individuals in our selection we do not have the statistical power<br>350 to explore concrete endpoints such as mortality or frailty in this study. We consider this<br>351 study as a proof of princ 351 study as a proof of principle design to explore the relevance of omics scores generated in the<br>352 ageing field in the context of additional omics levels, to better understand why a score 352 ageing field in the context of additional omics levels, to better understand why a score<br>353 predicts endpoints and by what parameters the predictive power may potentially be 353 predicts endpoints and by what parameters the predictive power may potentially be<br>354 improved.

predicts endpoints and by what parameters the predictive power may potentially be<br>354 improved.<br>1353 In conclusion, our study confirms MetaboHealth as a robust marker for the inflammatory<br>1356 aspects of aging. In accordan 355 In concl<br>356 aspects of 356 aspects of aging. In accordance with the current biological aging theories, this score<br>357 effectively identifies individuals exhibiting reduced lymphocyte counts and increased levels 357 effectively identifies individuals exhibiting reduced lymphocyte counts and increased levels<br>358 of pro-inflammatory proteins regardless of chronological aging. We believe that this<br>359 investigation supports integrati 358 of pro-inflammatory proteins regardless of chronological aging. We believe that this<br>359 investigation supports integrating the MetaboHealth score with proteome data to enhance 359 investigation supports integrating the MetaboHealth score with proteome data to enhance<br>360 the prognostic value of the score, increase our comprehension of the aging process and loss 360 the prognostic value of the score, increase our comprehension of the aging process and loss<br>361 of health in individuals identified by the score and the health gain that may be expected by<br>362 timely intervention. of health in individuals identified by the score and the health gain that may be expected by 362 timely intervention.<br>363<br>363  $363$ 

364 **MATERIALS AND METHODS**<br>365 **Study design**<br>366 This study was actuated starting fr 365 Study design<br>366 This study was<br>367 consortium origina This study was actuated starting from the metabolomics data included in the BBMRI-NL<br>367 consortium originated from the participants to three cohorts: Leiden Longevity Study (LLS-<br>368 PAROFFS), Rotterdam Study (RS), and th 1369 Longevity study is a population-based cohort with a unique two-generation design,<br>370 examining 421 Dutch long-lived families [37]. The current work was performed on the first 370 examining 421 Dutch long-lived families [37]. The current work was performed on the first<br>371 measurements (IOP1) of the second generation, namely the Offspring and their Partners, for 371 measurements (IOP1) of the second generation, namely the Offspring and their Partners, for<br>372 a total of 2,313 participants. The Rotterdam Study is a population-based prospective study<br>373 on individuals living in a s a total of 2,313 participants. The Rotterdam Study is a population-based prospective study 373 on individuals living in a specific neighborhood in Rotterdam, prone to cardiovascular<br>374 endpoints [38]. The current study was set off utilizing the first measurements (RS-I), which 374 endpoints [38]. The current study was set off utilizing the first measurements (RS-I), which<br>375 enclosed a total of 2,986 participants. The Netherlands Twin Register is a prospective study 375 enclosed a total of 2,986 participants. The Netherlands Twin Register is a prospective study<br>376 investigating young and adult twins and multiples along with their family members. In this 376 investigating young and adult twins and multiples along with their family members. In this<br>377 particular study we focused on the monozygotic dizygotic twin pairs within in the cohort, for 377 particular study we focused on the monozygotic dizygotic twin pairs within in the cohort, for<br>378 heritability estimation and implemented a within pairs case-control study design for the 25 377 particular study we focused on the monozygotic dizygotic twin pairs within in the cohort, for<br>378 heritability estimation and implemented a within pairs case-control study design for the 25<br>379 most discordant MZ twin 379 here in pairs estimation and increases a with the 25 most discordant MZ twin pairs.

### 381 **Data and sample collection**

### 382

382 Metabolomics measurements<br>383 The metabolomics dataset was 383 The metabolomics measurements<br>383 The metabolomics dataset was<br>384 plasma samples of the entire coh 384 plasma samples of the entire cohorts (LLS-PAROFFS: 2,313 samples, RS: 2,986 samples, and<br>385 NTR: samples). The features were quantified using the high-throughput proton Nuclear 385 NTR: samples). The features were quantified using the high-throughput proton Nuclear<br>386 Magnetic Resonance (<sup>1</sup>H-NMR) platform made available by Nightingale Health Ltd., Helsinki, 386 Magnetic Resonance (<sup>1</sup>H-NMR) platform made available by Nightingale Health Ltd., Helsinki, 386 Magnetic Resonance (\*H-NMR) platform made available by Nightingale Health Ltd., Helsinki,<br>,<br>,

387 388 Iipids, lipoproteins, fatty acid composition, various amino acids, and their derived measures<br>389 (e.g., ratios) [39,40]. We employed the dataset quantified in 2014 to ensure the correct 389 (e.g., ratios) [39,40]. We employed the dataset quantified in 2014 to ensure the correct<br>390 projection of the MetaboHealth model, originally trained on this version of the platform. 389 (e.g., ratios) [39,40]. We employed the dataset quantified in 2014 to ensure the cor<br>390 projection of the MetaboHealth model, originally trained on this version of the platform.<br>391 391 Selection of the participants from the large population studies

## 392<br>393

393 **LLS-PAROFFS and RS:** The sample selection was based of<br>394 independent part of the MetaboHealth, which was obtained as t 233 LLS-PAROFFS and RS: The sample selection was based on the chronological age<br>394 independent part of the MetaboHealth, which was obtained as the residual from a linear<br>395 regression of chronological age on the metabolo 394 independent part of the MetaboHealth, which was obtained as the residual from a linear<br>395 regression of chronological age on the metabolomics score. The cases were defined as the<br>396 25 participants with the highest M 25 participants with the highest MetaboHealth, indicating an unhealthier status, within each<br>397 cohort separately. Following, to limit the confounding effect of sex and age, the controls<br>398 were selected as the participa 397 cohort separately. Following, to limit the confounding effect of sex and age, the controls<br>398 were selected as the participants with the lowest score that could have at least one match<br>399 with the cases in terms of b were selected as the participants with the lowest score that could have at least one match<br>399 with the cases in terms of both age, and sex.<br>400 NTR: The twin population of NTR was composed on 726 complete monozygotic (MZT

300 **NTR:** The twin population of NTR was composed on 726 complete monozygotic (MZTs)<br>3988 and 450 complete dizygotic (DZTs) twin pairs. We extracted the 25 twin pairs which showed 399<br>400 **NTR:** The twin population of NTR was co<br>401 and 450 complete dizygotic (DZTs) twin pairs 400 **NTR:** The twin population of NTR was composed on 726 complete monozygotic (MZTs)<br>401 and 450 complete dizygotic (DZTs) twin pairs. We extracted the 25 twin pairs which showed<br>402 the largest differences in MetaboHealt 402 the largest differences in MetaboHealth score.<br>403

## 403  $\frac{3}{2}$  Cytokines quantification 404<br>405

Cytokines quantification<br>405 We used previously validated multiplex immunoassays (Luminex platform) to determine<br>406 plasma protein levels [41]. All assays were performed at the ISO-certified multiplex core<br>407 facility of 406 plasma protein levels [41]. All assays were performed at the ISO-certified multiplex core<br>407 facility of the UMC Utrecht, Utrecht, The Netherlands. Before analysis all samples were 407 facility of the UMC Utrecht, Utrecht, The Netherlands. Before analysis all samples were<br>408 centrifuged through 0.22  $\mu$ m spin-X filtration columns (Corning, Corning NY USA) to remove 408 centrifuged through 0.22 µm spin-X filtration columns (Corning, Corning NY USA) to remove<br>409 debris. Non-specific (heterophillic) antibodies, which may interfere with the assay, were debris. Non-specific (heterophillic) antibodies, which may interfere with the assay, were<br>410 blocked using Heteroblock (Omega Biologicals, Bozeman, MT, USA) as previously described 409 debris. Non-specific (heterophillic) antibodies, which may interfere with the assay, were<br>410 blocked using Heteroblock (Omega Biologicals, Bozeman, MT, USA) as previously described

 $\frac{1}{\sqrt{2}}$ 

 $412$  Sanquin, Amsterdam, the Netherlands). We determined levels of the immunoregulating<br>413 proteins IL-1β, IL-2, IL-6, IFN-γ, GDF-15, CCL2/MCP-1, CCL3/MIP-1α, CCL5/RANTES, 413 proteins IL-1 $\beta$ , IL-2, IL-6, IFN- $\gamma$ , GDF-15, CCL2/MCP-1, CCL3/MIP-1 $\alpha$ , CCL5/RANTES,<br>414 CCL11/Eotaxin, CXCL1/GRO-1 $\alpha$ , CXCL9/MIG, PAI-1, BDNF, TRAIN and soluble P-selectin in 414 CCL11/Eotaxin, CXCL1/GRO-1α, CXCL9/MIG, PAI-1, BDNF, TRAIN and soluble P-selectin in<br>415 plasma. These markers were selected based on two critera, first that they could be 415 plasma. These markers were selected based on two critera, first that they could be<br>416 measured on our budgetary boundaries using the Luminex platform in the Utrecht plasma. These markers were selected based on two critera, first that they could be<br>416 measured on our budgetary boundaries using the Luminex platform in the Utrecht<br>417 University; secondly because they were previously in 417 University; secondly because they were previously indicated as markers of ageing, mortality,<br>418 frailty and age-related disease. The last criterium was in great part based on the following 418 frailty and age-related disease. The last criterium was in great part based on the following<br>419 studies: A) belonging to the top 15 most informative variables in the iAge clock [25] (for the 419 studies: A) belonging to the top 15 most informative variables in the iAge clock [25] (for the<br>420 cytokines: CXCL9/MIG, EOTAXIN, CCL3/Mip-1 $\alpha$ , IL-1 $\beta$ , IFN- $\gamma$ , CXCL5/RANTES;,CXCL1/GRO-1 $\alpha$ , 420 cytokines: CXCL9/MIG, EOTAXIN, CCL3/Mip-1α, IL-1β, IFN-γ, CXCL5/RANTES;,CXCL1/GRO-1α,<br>421 CCL2/MCP1, IL-2; TRAIL, PAI-1); B) recognized as indicators of frailty (Pselectin and BNDF) 421 CCL2/MCP1, IL-2; TRAIL, PAI-1); B) recognized as indicators of frailty (Pselectin and BNDF)<br>422 [26]; C) or marker of chronic inflammation (IL6) [27], and d) finally, widely explored marker 422 [26]; C) or marker of chronic inflammation (IL6) [27], and d) finally, widely explored marker<br>423 for aging, cancer, cardiovascular, and lung disease (GDF-15/MIC-1) [28]. for aging, cancer, cardiovascular, and lung disease (GDF-15/MIC-1) [28].<br>424 Although the majority of the features (73%) were quantified in undiluted samples, PAI1<br>425 was measured with a dilution rate of 1/10, and, GDF15,

424 Although the majority of the features (73%) were quantified in und<br>425 was measured with a dilution rate of 1/10, and, GDF15, RANTES, and B was measured with a dilution rate of  $1/10$ , and, GDF15, RANTES, and BDNF, with a dilution rate of  $1/100$ . 426 was measured with a dilution rate of  $1/100$ .<br>427

## 427<br>428 **Protein dig** 427<br>428<br>429

428 Protein digestion<br>429 Protein digestion<br>430 Vollmy et. al. [18]. A Protein digestion was performed on an Agilent Bravo liquid handling Platform following<br>430 Vollmy et. al. [18]. A pooled QC sample was prepared by pooling an equal amount of each<br>431 digested sample. Then all samples were 431 digested sample. Then all samples were diluted 40 times with TFA 1% to bring them to an<br>432 approximate concentration of 10 ng/uL. Finally, the samples were loaded onto Evotips 432 approximate concentration of 10 ng/uL. Finally, the samples were loaded onto Evotips<br>433 (Odense, Danemark) using an Agilent Bravo liquid handling platform. 433 (Odense, Danemark) using an Agilent Bravo liquid handling platform.<br>434

 $434$ 

435 **LC-MS data acquisition**<br>436 The digested samples w<br>437 Danemark) on a EV-1109 c 437 Danemark) on a EV-1109 column (C18, 8 cm x 150 μm, 1.5 μm, Evosep, Danemark) coupled<br>438 to a timsTOF-HT (Bruker, Germany) equipped with a Captive Spray source and operating in 438 to a timsTOF-HT (Bruker, Germany) equipped with a Captive Spray source and operating in<br>439 DIA-PASEF adapted from Skowronek et. al. [42] Briefly, two ion mobility windows per dia-1439 DIA-PASEF adapted from Skowronek et. al. [42] Briefly, two ion mobility windows per dia-<br>140 PASEF scan with 12 variable isolation window widths adjusted to the precursor densities 440 PASEF scan with 12 variable isolation window widths adjusted to the precursor densities<br>441 were used. The ion mobility range was set to 0.6 and 1.5 cm<sup>-1</sup>. The accumulation and ramp<br>442 times were specified as 100 ms were specified as 100 ms for all experiments. Source capillary voltage and<br>143 temperature were set to 1800 V and 180°C. Drving and sheath gas were set to 3 L/min. The 443 temperature were set to 1800 V and 180°C. Drying and sheath gas were set to 3 L/min. The<br>444 pooled QC samples were injected every 8 samples. 1443 temperature were set to 1800 V and 1800 V and sheath gas were set to 3 L/min. The<br>444 pooled QC samples were injected every 8 samples.

## 445<br>446 Processing of proteomics data 445<br>446<br>447

Processing of proteomics data<br>
Raw data were processed using DIA-NN 1.8.1 [43], Peptides were searched against an in-<br>
silico predicted library computed from the human proteome with isoforms (UniProtKB and<br>
449 TrEMBL, 103 TreMBL, 103830 protein entries and 20560 genes) and the common protein contaminants<br>450 with 2 missed-cleavages and no variable modification. Match Between Run was used and<br>451 the Heuristic inference was disabled. MS1 and with 2 missed-cleavages and no variable modification. Match Between Run was used and<br>451 the Heuristic inference was disabled. MS1 and MSMS mass accuracy were set to 10 and 20<br>452 ppm respectively. 451 the Heuristic inference was disabled. MS1 and MSMS mass accuracy were set to 10 and 20<br>452 ppm respectively.<br>453 The protein intensities were computed using the maxLFQ algorithm implemented in the

The protein intensities were computed using the maxLFQ algorithm implemented in the<br>454 DIA-NN R-package. For this only the prototypic precursors that satisfy the following criteria 453 The protein int<br>454 DIA-NN R-package 454 DIA-NN R-package. For this only the prototypic precursors that satisfy the following criteria<br>455 were considered: Q.value ≤ 1%, missing value ≤ 20% and RSD ≤ 40% in the QC samples. The<br>456 protein groups were filter 455 were considered: Q.value ≤ 1%, missing value ≤ 20% and RSD ≤ 40% in the QC samples. The 456 protein groups were filtered at lib.Q.Value ≤ 1%, lib.PG.Q.value≤ 1%.<br>457 456 were considered: Q. value  $\geq$  1%, missing value  $\geq$  1% and  $R\geq$  2004 and RSD  $\geq$  1%.<br>457 457 protein groups were filtered at library were  $\frac{1}{2}$ .

458 **Covariates**<br>459 Data on age (<br>460 and blood press Data on age (in years), sex (males/females), BMI (kg/m<sup>2</sup>), cell counts (%), lipid medication,<br>460 and blood pressure lowering medication were reported within the BBMRI-nl Consortium.<br>461 These covariates were evaluated as 461 These covariates were evaluated as possible confounders as they are known to be associated<br>462 with both the metabolomics and the proteomics datasets. Age and sex were self-reported, 462 with both the metabolomics and the proteomics datasets. Age and sex were self-reported,<br>463 and BMI was calculated based on weight and height. The cell counts percentage was defined 463 and BMI was calculated based on weight and height. The cell counts percentage was defined<br>464 as the measured monocytes and lymphocytes percentage, taking the granulocyte and BMI was calculated based on weight and height. The cell counts percentage was defined<br>464 as the measured monocytes and lymphocytes percentage, taking the granulocyte<br>465 percentage as given.<br>466 465 percentage as given.<br>466

## **Statistical Analyses** 467<br>468

Abriga Preprocessing<br>169 Metabolomics and MetaboHealth score projection: We applied the same quality control 468 Preprocessing<br>469 Metabolomics<br>470 previously descril 469 Metabolomics and Metaboriean missive projection: We applied the same quality control<br>470 previously described by Deelen et al. [11], using the R package MiMIR [44]. While the<br>471 Nightingale Health platform measures ov 471 Nightingale Health platform measures over 250 metabolomics features, we focused our<br>472 attention on the 14 metabolomics variables included in the MetaboHealth model (list of According 1471 Nightingale Health platform measures over 250 metabolomics features, we focused our<br>472 attention on the 14 metabolomics variables included in the MetaboHealth model (list of<br>473 analytes can be found in the 473 analytes can be found in the Supplementary Table S1). Then, we applied a logarithm<br>474 transformation to the analytes, while adding a value of 1 to all analytes containing any zero 474 transformation to the analytes, while adding a value of 1 to all analytes containing any zero<br>475 as a value. Afterwards a z-scale normalization was applied separately within each cohort to 475 as a value. Afterwards a z-scale normalization was applied separately within each cohort to<br>476 minimize batch effects. Finally, we projected the MetaboHealth score using the coefficients<br>477 indicated by Deelen et al. minimize batch effects. Finally, we projected the MetaboHealth score using the coefficients<br>477 indicated by Deelen et al. [11,44].<br>478 Cytokines: Initially, we assessed the occurrence of samples reported as under the lowe

478 Cytokines: Initially, we assessed the occurrence of samples reported as under the lower<br>479 detection thresholds, comparing cases and controls, separately for each feature. 478 Cytokines: Initially, we assessed<br>479 detection thresholds, comparin 478 Cytokines: Initially, we assessed the occurrence of samples reported as under the lower<br>479 detection thresholds, comparing cases and controls, separately for each feature.<br>480 Subsequently, we employed the Fisher test 480 Subsequently, we employed the Fisher test to determine the statistical significance of these  $\frac{1}{\sqrt{2}}$  subsequently, we employ the statistical significance of the statistical significance of these statistical significance of the statistical significance of these statistical significance of these statistical si

482 that consistently fell<br>483 (namely, IL2, TRAIL,  $\alpha$ difference (p-value $\leq$  0.05). Following, we proceeded to exclude six out of fifteen features<br>that consistently fell below the lower detection thresholds in most of the quantified samples<br>(namely, IL2, TRAIL, GRO1a, IFNg that consistently fell below the lower detection thresholds in most of the quantified samples<br>483 (namely, IL2, TRAIL, GRO1a, IFNg, IL1b, and PAI1 were undetected in more than 65% of the<br>484 samples). There were no values 484 samples). There were no values recognized as outlier samples, considered as values more<br>485 than 5 standard deviations (SD) away from the mean of the feature. Finally, the remaining 485 than 5 standard deviations (SD) away from the mean of the feature. Finally, the remaining<br>486 nine features were first log transformed and then z scaled separately per biobank to reduce than 5 standard deviations (SD) away from the mean of the feature. Finally, the remaining<br>186 in the features were first log transformed and then z scaled separately per biobank to reduce<br>187 batch effects.

488 Proteomics: We applied a similar approach to the proteomics. We evaluated the 488 *Proteomics*<br>489 differential pa 488 Proteomics: We applied a similar approach to the proteomics: We evaluated the<br>489 differential patterns in missing values of the features between the cases and controls.<br>490 Features with more than 5% missing values (2 490 Features with more than 5% missing values (20 missing) were subsequently excluded,<br>491 resulting in a set of 261 analytes, from the 337 initial set 8 values (0.03%) were recognized as 491 Features with more than 5 of the set of 261 analytes, from the 337 initial set 8 values (0.03%) were recognized as<br>492 Features wearing that they resulted as the values 5 SD away from their mean of the feature 492 outliers, meaning that they resulted as the values 5 SD away from their mean of the feature<br>493 and set as missing information. We imputed the remaining 114 missing values (0.4%) using outliers, meaning that they resulted as the values 5 SD away from their mean of the feature<br>and set as missing information. We imputed the remaining 114 missing values (0.4%) using<br>the non-linear iterative partial least sq 494 the non-linear iterative partial least squares method (nipals), implemented in the R package<br>495 *pcaMethods*. Finally, to enhance comparability and facilitate downstream analysis we 495 pcaMethods. Finally, to enhance comparability and facilitate downstream analysis we<br>496 performed a log transformation and a z-scaling of the features per biobank. 495 pcawethods. Finally, to enhance comparability and facilitate downstream analysis we<br>496 performed a log transformation and a z-scaling of the features per biobank.<br>497 496 performed a log transformation and a z-scaling of the features per biobank.<br>497 **Linear Association and Meta-analyses in RS and LLS-PAROFFS** 

## 498<br>499

199 Linear Association and Meta-analyses in RS and LLS-PAROFFS<br>499 Initially, we applied linear regression models across the en<br>500 associations between each cytokines and proteomics featu 499 Initially, we applied linear regression models across the entire activative in across the entire dataset of the entire and proteomics features separately, with the case/control status of the participants. These analyse 501 case/control status of the participants. These analyses accounted for potential confounding<br>502 factors (age, sex, BMI, cell counts, and usage of lipid medication and blood pressure 502 factors (age, sex, BMI, cell counts, and usage of lipid medication and blood pressure<br>503 lowering medication). Subsequently, we attempted to evaluate the associations 503 lowering medication). Subsequently, we attempted to evaluate the associations<br>504 independently for each cohort (LLS-PAROFFS and RS) and conducted a meta-analysis. The 504 independently for each cohort (LLS-PAROFFS and RS) and conducted a meta-analysis. The subsequently for each cohort (LLS-PAROFFS and RS) and conducted a meta-analysis. The 504 independently for each cohort (LLS-PAROFFS and RS) and conducted a meta-analysis. The

505 meta-analysis was performed with a restricted maximum likelihood estimator using the<br>506 package *metafor* in R. To correct for multiple testing, we applied False Discovery Rate (FDR).<br>507 506 package *metafor* in R. To correct for multiple testing, we applied False Discovery Rate (FDR).<br>507<br>508 **Enrichment Analyses and Network analysis** 

## 508<br>509

509 Enrichment Analyses and Network analysis<br>509 We performed the functional enrichment<br>510 using the web-tool *Enrichr* [45]. We evaluated 1199 We performed the functional environment of the most interesting processive features<br>510 We using the web-tool Enrichment [45]. We evaluated the Gene Ontology (GO) Biological Processes<br>511 (BP), the KEGG pathways, and using the web-tool *Enrichr* [45]. We evaluated the Gene Ontology (GO) Biological Processes<br>511 (BP), the KEGG pathways, and Reactome. Firstly, we performed an enrichment analyses on<br>512 the full set of 337 proteomics feat 512 the full set of 337 proteomics features to evaluate the functions overall characterizing to the<br>513 proteomics platform (Figure S7A-B). Secondly, we analyzed separately the 38 positively and<br>514 the 68 negatively assoc 5<br>513 proteomics platform (Figure S7A-B). Secondly, we analyzed separately the 38 positively and<br>514 the 68 negatively associated proteins to the MetaboHealth's extremes corrected for sex, age, 514 the 68 negatively associated proteins to the MetaboHealth's extremes corrected for sex, age,<br>515 and BMI (Figure S7C-H). To ensure a fair enrichment analysis of the significantly associated 515 and BMI (Figure S7C-H). To ensure a fair enrichment analysis of the significantly associated<br>516 proteins we used the full list of 337 proteins as the background of possible analytes. To 516 proteins we used the full list of 337 proteins as the background of possible analytes. To<br>517 better interpret at the GO BP results we used the R package rrvgo (threshold=0.7), which is 516 proteins we used the full list of 337 proteins as the background of possible analytes. To<br>517 better interpret at the GO BP results we used the R package *rrvgo* (threshold=0.7), which is<br>518 able to summarize the redu 517 better interpret at the GO BP results we used the R package *RVgo* (threshold=0.7), which is<br>518 able to summarize the redundant information in the database [46].

## 519<br>520 Analyses in the Netherlands Twin Register 520<br>521

520 **Analyses in the Netherlands Twin Register**<br>521 First, we examined heritability of the N<br>522 Register. Monozygotic twins (MZTs) share th 522 Register. Monozygotic twins (MZTs) share their (almost) complete DNA sequence, while<br>523 dizygotic twins (DZTs) share on average 50% of their segregating genes. Any differences in S22 Register. Monozygotic twins (MZTs) share their (almost) complete DNA sequence, while<br>
523 dizygotic twins (DZTs) share on average 50% of their segregating genes. Any differences in<br>
524 correlations between MZTs and b 524 correlations between MZTs and between DZTs offers a first hint on genetic influences in the<br>525 signal. We obtained an estimate of the heritability of a trait  $(h^2)$  using the formula:  $h^2 =$ 525 signal. We obtained an estimate of the heritability of a trait (h<sup>2</sup>) using the formula: h<sup>2</sup> =<br>526 2(r<sub>MZTs</sub>-r<sub>DZTs</sub>), where r denotes the correlation between the twins [19]. S25 signal. We obtained an estimate of the heritability of a trait (h<sup>2</sup>) using the formula: h<sup>2</sup> =<br>526 = 2( $r_{MZTs}$ -r<sub>DZTs</sub>), where r denotes the correlation between the twins [19].<br>527 = Secondly, we assessed the associ

Secondly, we assessed the associations between the cytokine<br>528 features with the case/control status of the selected MZTs with the 528 features with the case/control status of the selected MZTs with the highest MetaboHealth  $528.8 \pm 0.000$  features with the selected MZTs with the selected MZTs with the highest MZTs with the highest M $\alpha$ 

- 529 529 differences. To do this we used linear mixed models to consider the family status, while<br>530 allowing to systematically correct for potential covariates. Finally, FDR to correct for multiple<br>531 testing.
- 531 desting.<br>532 allowing to systematically correct for potential covariates. Finally, FDR to correct for multiple<br>532
- 
- 

## 532<br>532<br>533 **Dat**: 533<br>534

- 533 **Data sharing**<br>534 Mass Spectrometr<br>535 via the PRIDE parti Mass Spectrometry data have been deposited to the ProteomeXchange Consortium<br>1535 via the PRIDE partner repository with the dataset identifier PXD057946. Phenotypic<br>1536 information are available upon reasonable request at
- via the PRIDE partner repository with the dataset identifier PXD057946. Phenotypic<br>536 information are available upon reasonable request at BBMRI-nl<br>537 https://www.bbmri.nl/services/samples-images-data.
- 536 information are available upon reasonable request at BBMRI-nl<br>537 https://www.bbmri.nl/services/samples-images-data.<br>538 **Acknowledgements**
- 

# For the control of the control of the services of the ARMRI M.<br>1986 – This work was nerformed within the BBMRI M.

538 **Acknowledgements**<br>539 This work was performed<br>540 BBMRI-NL (financed by NWO 540 BBMRI-NL (financed by NWO 184.021.007 and 184.033.111), X-omics (NWO 184.034.019),<br>541 VOILA (ZonMW 457001001) and Medical Delta (METABODELTA: Metabolomics for clinical 541 VOILA (ZonMW 457001001) and Medical Delta (METABODELTA: Metabolomics for clinical<br>542 advances in the Medical Delta). EvdA is funded by a personal grant of the Dutch Research<br>543 Council (NWO: VENI:09150161810095). Ack 542 advances in the Medical Delta). EvdA is funded by a personal grant of the Dutch Research<br>543 Council (NWO; VENI:09150161810095). Acknowledgements for all contributing studies can 542 advances in the Medical Delta). EvdA is funded by a personal grant of the Dutch Research<br>543 Council (NWO; VENI:09150161810095). Acknowledgements for all contributing studies can<br>544 be found in the Supplementary Mater 544 be found in the Supplementary Material-BIOS Consortium. Finally, we would like to thank<br>545 Prof. Peter Bram 't Hoen for critically appraising this manuscript. 545 before the Supplementary Material-BIOS Contributions (1944), we would like to thank<br>545 Prof. Peter Bram 't Hoen for critically appraising this manuscript.<br>546 **Contributors** 

546 **Contributors**<br>547 PES, DB, EbvDA, and MJTR conceived and wrote the manuscript. DB performed the 546 **Contributors**<br>547 PES, DB, EbvDA,<br>548 analyses. EBvDA and<br>549 EbvDA, and MJTR we 548 analyses. EBvDA and MJTR verified and supervised the analyses. PES, MB, DiB, AjvG, DB,<br>549 EbvDA, and MJTR were involved in defining the study design. PES, MB, and NL were involved 549 EbvDA, and MJTR were involved in defining the study design. PES, MB, and NL were involved<br>550 in the data acquisition for the LLS-PAROFF cohort; RP and DiB in the data acquisition of the 550 in the data acquisition for the LLS-PAROFF cohort; RP and DiB in the data acquisition of the<br>551 Netherland Twin Register; JmV and MG in the data acquisition of the Rotterdam Study. AjrH, 550 in the data acquisition for the LLS-PAROFF cohort; RP and DiB in the data acquisition of the<br>551 Netherland Twin Register; JmV and MG in the data acquisition of the Rotterdam Study. AjrH, 551 Netherland Twin Register; JmV and MG in the data acquisition of the Rotterdam Study. AjrH,

- 552
- 552 ND, and KeS performed the Mass Spectrometry proteomics. EFK performed the cytokine<br>quantification. All authors discussed the results and contributed to the final manuscript.<br>554 **Competing Interests**

554 **Competing Interests**<br>555 Authors declare no competing interests. 555 Authors declare no competing interests.<br>556 **REFERENCES**<br>557 1. Partridge. L.: Deelen. J.: Slagboom. P.

- 557 1. Partridge, L.; Deelen, J.; Slagboom, P.E. Facing up to the Global Challenges of Ageing.<br>558 *Nature* **2018**, 561, 45–56, doi:10.1038/s41586-018-0457-8.
- 556 **REFERENCES**<br>557 1. Partridge, L.; De<br>558 *Nature* **2018**, 56<br>559 2. Comfort, A. TES<br>560 1411–1415, doi<br>561 3. Moqri, M.; Herz
- 558 Mature 2018, 561, 45–56, doi:10.1038/s41586-018-0457-8.<br>559 2. Comfort, A. TEST-BATTERY TO MEASURE AGEING-RATE IN MAN. The Lancet 1969, 294,<br>560 1411–1415, doi:10.1016/S0140-6736(69)90950-7.<br>561 3. Moqri, M.; Herzog, C Mature 2018, 501, 45–50, doi:10.1038/341588-018-0457-8.<br>559 2. Comfort, A. TEST-BATTERY TO MEASURE AGEING-RATE IN N<br>560 1411–1415, doi:10.1016/S0140-6736(69)90950-7.<br>561 3. Moqri, M.; Herzog, C.; Poganik, J.R.; Biomarkers 559 2. Comfort, A. TEST-BATTERY TO MEASURE AGEING-RATE IN MAN. The Eancet 1969, 294,<br>560 1411–1415, doi:10.1016/S0140-6736(69)90950-7.<br>561 3. Moqri, M.; Herzog, C.; Poganik, J.R.; Biomarkers of Aging Consortium; Justice, J 561 3. Moqri, M.; Herzog, C.; Poganik, J.R.; Biomarkers of<br>562 D.W.; Higgins-Chen, A.; Moskalev, A.; Fuellen, G.; Co.<br>563 for the Identification and Evaluation of Longevity<br>564 3775, doi:10.1016/j.cell.2023.08.003.<br>565 4. 161 562 D.W.; Higgins-Chen, A.; Moskalev, A.; Fuellen, G.; Cohen, A.A.; et al. Biomarkers of Aging<br>563 for the Identification and Evaluation of Longevity Interventions. *Cell* 2023, 186, 3758–<br>564 3775, doi:10.1016/j.cell.
- 563 for the Identification and Evaluation of Longevity Interventions. *Cell* 2023, 186, 3758–<br>564 3775, doi:10.1016/j.cell.2023.08.003.<br>565 4. van den Akker Erik B.; Trompet Stella; Barkey Wolf Jurriaan J.H.; Beekman Maria 563 for the Identification and Evaluation of Longevity Interventions. Cell 2023, 186, 3738–<br>564 3775, doi:10.1016/j.cell.2023.08.003.<br>565 4. van den Akker Erik B.; Trompet Stella; Barkey Wolf Jurriaan J.H.; Beekman Marian; 565 4. van den Akker Erik B.; Trompet Stel<br>566 Suchiman H. Eka D.; Deelen Joris; Asse<br>567 Petra M.; et al. Metabolic Age Ba<br>568 Repository as Biomarker of Age-Rela<br>569 *Medicine* 2020, 13, 541–547, doi:10.11 566 567 5. Suchiman H. Eka D.; Deelen Joris; Asselbergs Folkert W.; Boersma Eric; Cats Davy; Elders<br>567 Petra M.; et al. Metabolic Age Based on the BBMRI-NL 1H-NMR Metabolomics<br>568 Repository as Biomarker of Age-Related Di
- Repository as Biomarker of Age-Related Disease. Circulation: Genomic and Precision<br>569 Medicine 2020, 13, 541–547, doi:10.1161/CIRCGEN.119.002610.<br>570 5. Group, N.H.B.C.; Barrett, J.C.; Esko, T.; Fischer, K.; Jostins-Dean, For Petra May as Biomarker of Age-Related Disease. *Circulation: Genomic and Precision*<br>569 Medicine 2020, 13, 541–547, doi:10.1161/CIRCGEN.119.002610.<br>570 5. Group, N.H.B.C.; Barrett, J.C.; Esko, T.; Fischer, K.; Jostins-Repository as Biomarker of Age-Related Disease. Circulation: Genomic dila Trecision<br>569 Medicine 2020, 13, 541–547, doi:10.1161/CIRCGEN.119.002610.<br>570 5. Group, N.H.B.C.; Barrett, J.C.; Esko, T.; Fischer, K.; Jostins-Dean Sob Medicine 2020, 13, 341–547, doi:10.1161, CircoEN.119.002010.<br>570 5. Group, N.H.B.C.; Barrett, J.C.; Esko, T.; Fischer, K.; Jostins-D<br>571 Julkunen, H.; Jääskeläinen, T.; Kerimov, N.; Kerminen, S.; et<br>572 Genomic Predict Julkunen, H.; Jääskeläinen, T.; Kerimov, N.; Kerminen, S.; et al. Metabolomic and<br>572 Genomic Prediction of Common Diseases in 477,706 Participants in Three National<br>573 Biobanks 2023, 2023.06.09.23291213.<br>574 G. Yu, B.; Z
- 572 Genomic Prediction of Common Diseases in 477,706 Participants in Three National<br>573 Biobanks 2023, 2023.06.09.23291213.<br>574 6. Yu, B.; Zanetti, K.A.; Temprosa, M.; Albanes, D.; Appel, N.; Barrera, C.B.; Ben-Shlomo, Y.; 573 Biobanks 2023, 2023.06.09.23291213.<br>574 6. Yu, B.; Zanetti, K.A.; Temprosa, M.; Albanes, D.; Appel, N.; Barrera, C.B.; Ben-Shlomo, Y.;<br>575 Boerwinkle, E.; Casas, J.P.; Clish, C.; et al. The Consortium of Metabolomics S 574 6. Yu, B.; Zanetti, K.A.; Temprosa, M.; Albanks 2023.<br>575 Boerwinkle, E.; Casas, J.P.; Clish, C.; e<br>576 (COMETS): Metabolomics in 47 Prospe<br>577 991–1012, doi:10.1093/aje/kwz028.<br>578 7. Nightingale Health UK Biobank Ini 575 Boerwinkle, E.; Casas, J.P.; Clish, C.; et al. The Consortium of Metabolomics Studies<br>576 (COMETS): Metabolomics in 47 Prospective Cohort Studies. Am J Epidemiol 2019, 188,<br>577 991–1012, doi:10.1093/aje/kwz028.<br>578 7.
- (COMETS): Metabolomics in 47 Prospective Cohort Studies. Am J Epidemiol 2019, 188,<br>577 991–1012, doi:10.1093/aje/kwz028.<br>578 7. Nightingale Health UK Biobank Initiative; Julkunen, H.; Cichońska, A.; Slagboom, P.E.;<br>579 Wür 576 (COMETS): Metabolomics in 47 Prospective Conort Studies. Am J Epidemiol 2015, 188,<br>578 7. Nightingale Health UK Biobank Initiative; Julkunen, H.; Cichońska, A.; Slagboom, P.E.;<br>579 Würtz, P. Metabolic Biomarker Profili 578 7. Nightingale Health UK Biobank Initia<br>579 Würtz, P. Metabolic Biomarker Prof<br>580 Pneumonia and COVID-19 in the<br>581 doi:10.7554/eLife.63033.<br>582 8. 't Hart, L.M.; Vogelzangs, N.; Mook Würtz, P. Metabolic Biomarker Profiling for Identification of Susceptibility to Severe<br>580 Pneumonia and COVID-19 in the General Population. *eLife* 2021, 10, e63033,<br>581 doi:10.7554/eLife.63033.<br>582 8. 't Hart, L.M.; Voge
- 580 Pneumonia and COVID-19 in the General Population. *eLife* 2021, 10, e63033,<br>581 doi:10.7554/eLife.63033.<br>582 8. 't Hart, L.M.; Vogelzangs, N.; Mook-Kanamori, D.O.; Brahimaj, A.; Nano, J.; van der<br>583 Heijden, A.A.W.A.; Friedmonia and COVID-19 in the General Population. ELife 2021, 10, e63033,<br>581 8. († Hart, L.M.; Vogelzangs, N.; Mook-Kanamori, D.O.; Brahimaj, A.; Nano, J.; van der<br>583 Heijden, A.A.W.A.; Willems van Dijk, K.; Slieker, R. 582 8. 't Hart, L.M.; Vogelzangs<br>583 Heijden, A.A.W.A.; Willem<br>584 Blood Metabolomic Meas<br>585 Type 2 Diabetes. *The Jour*<br>586 4579, doi:10.1210/jc.2018 583 Heijden, A.A.W.A.; Willems van Dijk, K.; Slieker, R.C.; Steyerberg, E.W.; Ikram, M.A.; et al.<br>584 Blood Metabolomic Measures Associate With Present and Future Glycemic Control in<br>585 Type 2 Diabetes. *The Journal of Cl* 584 Blood Metabolomic Measures Associate With Present and Future Glycemic Control in<br>585 Type 2 Diabetes. *The Journal of Clinical Endocrinology & Metabolism* 2018, 103, 4569–<br>587 9. Bizzarri, D.; Reinders, M.J.T.; Beekman
- 584 Blood Metabolomic Measures Associate With Present and Future Glycemic Control in Fype 2 Diabetes. The Journal of Clinical Endocrmology & Metabolism 2018, 103, 4303<br>
586 Type 2 Diabetes. M.J.T.; Beekman, M.; Slagboom, P.E.; Bbmri-, N.; van den Akker,<br>
587 Bizzarri, D.; Reinders, M.J.T.; Beekman, M.; Sla 587 9. Bizzarri, D.; Reinders, M.J.T.; Beek<br>588 E.B. 1H-NMR Metabolomics-Based<br>589 and Endpoints. *EBioMedicine* 2022<br>590 10. Buergel, T.; Steinfeldt, J.; Ruyoga, C<br>591 Belzen, J.; Loock, L.; Kittner, P.; O
- E.B. 1H-NMR Metabolomics-Based Surrogates to Impute Common Clinical Risk Factors<br>589 and Endpoints. *EBioMedicine* 2022, 75, 103764, doi:10.1016/j.ebiom.2021.103764.<br>590 10. Buergel, T.; Steinfeldt, J.; Ruyoga, G.; Pietzne and Endpoints. *EBioMedicine* **2022**, 75, 103764, doi:10.1016/j.ebiom.2021.103764.<br>590 10. Buergel, T.; Steinfeldt, J.; Ruyoga, G.; Pietzner, M.; Bizzarri, D.; Vojinovic, D.; Upmeier zu<br>591 Belzen, J.; Loock, L.; Kittner, 589 and Endpoints. EbioMedicine 2022, 75, 103764, doi:10.1016/j.ebiom.2021.103764.<br>590 10. Buergel, T.; Steinfeldt, J.; Ruyoga, G.; Pietzner, M.; Bizzarri, D.; Vojinovic, D.; Upmeie<br>692 Individual Multidisease Outcomes. Na 590 10. Buergel, T.; Steinfeldt, J.; Ruyoga, G.; Pietzner, M.; Bizzarri, D.; Vojinovic, D.; Upmeier zu<br>591 Belzen, J.; Loock, L.; Kittner, P.; Christmann, L.; et al. Metabolomic Profiles Predict<br>592 Individual Multidisease Individual Multidisease Outcomes. Nat Med 2022, 28, 2309-2320, doi:10.1038/s41591-592 Individual Multidisease Outcomes. Nat Med 2022, 28, 2309–2320, doi:10.1038/s41591-

595 Boyd, A.; Zierer, J.; van den Akker, E.B.; Ala-Korpela, M.; et al. A Metabolic Profile of All-<br>596 Cause Mortality Risk Identified in an Observational Study of 44,168 Individuals. Nature<br>597 Communications 2019, 10, 1–

- 600 Metabolomic Biomarkers for Biological Age: A Comparative Analysis of Mortality and Cause Mortality Risk Identified in an Observational Study of 44,168 Individuals. Nuture<br>
Communications 2019, 10, 1–8, doi:10.1038/s41467-019-11311-9.<br>
12. Kuiper, L.M.; Polinder-Bos, H.A.; Bizzarri, D.; Vojinovic, D.; Val Eommunications 2013, 10, 1–8, doi:10.1038/s41407-013-11311-5.<br>598 12. Kuiper, L.M.; Polinder-Bos, H.A.; Bizzarri, D.; Vojinovic, D.; Vallerga<br>599 Dollé, M.E.T.; Ghanbari, M.; Voortman, T.; Reinders, M.J.T.; et<br>600 Metabolo 599 Dollé, M.E.T.; Ghanbari, M.; Voortman, T.; Reinders, M.J.T.; et al. Epigenetic and<br>600 Metabolomic Biomarkers for Biological Age: A Comparative Analysis of Mortality and<br>601 Frailty Risk. *The Journals of Gerontology:* 500 Metabolomic Biomarkers for Biological Age: A Comparative Analysis of Mortality and<br>501 Frailty Risk. *The Journals of Gerontology: Series A* **2023**, glad137,<br>502 doi:10.1093/gerona/glad137.<br>503 13. Zonneveld, M.H.; Kuh
- Frailty Risk. The Journals of Gerontology: Series A 2023, glad137,<br>602 doi:10.1093/gerona/glad137.<br>603 13. Zonneveld, M.H.; Kuhaili, N.A.; Mooijaart, S.P.; Slagboom, P.E.; Jukema, J.W.; Noordam,<br>604 R.; Trompet, S. Increas Frailty Risk. The Journals of Gerontology: Series A 2023, glad137,<br>602 doi:10.1093/gerona/glad137.<br>603 13. Zonneveld, M.H.; Kuhaili, N.A.; Mooijaart, S.P.; Slagboom, P.E.; Jukema, J.W.; Noordam,<br>604 R.; Trompet, S. Increas 13. Ionneveld, M.H.; Kuhaili, N.A<br>604 R.; Trompet, S. Increased 1H-<br>605 Declined Cognitive Performar<br>606 of Cardiovascular Disease 202<br>607 14. Gonggrijp, B.M.A.; van de We 804 R.; Trompet, S. Increased 1H-NMR Metabolomics-Based Health Score Associates with<br>605 Declined Cognitive Performance and Functional Independence in Older Adults at Risk<br>606 of Cardiovascular Disease 2023, 2023.12.21.233
- Example 19 Declined Cognitive Performance and Functional Independence in Older Adults at Risk<br>
606 of Cardiovascular Disease 2023, 2023.12.21.23300037.<br>
607 14. Gonggrijp, B.M.A.; van de Weijer, S.G.A.; Bijleveld, C.C.J.H. 606 of Cardiovascular Disease 2023, 2023.12.21.23300037.<br>607 14. Gonggrijp, B.M.A.; van de Weijer, S.G.A.; Bijleveld, C.C.J.H.; van Dongen, J.; Boomsma,<br>608 D.I. The Co-Twin Control Design: Implementation and Methodologica 14. Gonggrijp, B.M.A.; van de Weijer, S.G.A.; Bijleveld, C.C<br>608 D.I. The Co-Twin Control Design: Implementation and<br>609 Twin Res Hum Genet 2023, 1–8, doi:10.1017/thg.2023.3<br>610 15. Ernster, V.L. Nested Case-Control Studie 608 D.I. The Co-Twin Control Design: Implementation and Methodological Considerations.<br>609 Twin Res Hum Genet 2023, 1–8, doi:10.1017/thg.2023.35.<br>610 15. Ernster, V.L. Nested Case-Control Studies. Preventive Medicine 1994,
- 
- 7Win Res Hum Genet 2023, 1–8, doi:10.1017/thg.2023.35.<br>610 15. Ernster, V.L. Nested Case-Control Studies. Preventive Medicine 1994, 23, 587–590,<br>611 doi:10.1006/pmed.1994.1093.<br>612 16. Bizzarri, D.; Reinders, M.J.T.; Kuipe Fram Genet 2023, 1–8, doi:10.1017/thg.2023.33.<br>610 15. Ernster, V.L. Nested Case-Control Studies. Preventive M<br>611 doi:10.1006/pmed.1994.1093.<br>612 16. Bizzarri, D.; Reinders, M.J.T.; Kuiper, L.M.; Beekman, M.;<br>613 Dongen, 610 15. Ernster, V.L. Nested Case-Control Studies. Preventive Medicine 1994, 23, 587–590,<br>611 doi:10.1006/pmed.1994.1093.<br>613 Dongen, J. van; Pool, R.; Boomsma, D.I.; Ghanbari, M.; et al. 1H-NMR Metabolomics-<br>614 Guided DN 612 16. Bizzarri, D.; Reinders, M.J.T.; K<br>613 Dongen, J. van; Pool, R.; Boom<br>614 Guided DNA Methylation Morta<br>615 17. Geyer, P.E.; Kulak, N.A.; Pichler,<br>616 Profiling to Assess Human
- Ether Bongen, J. van; Pool, R.; Boomsma, D.I.; Ghanbari, M.; et al. 1H-NMR Metabolomics-<br>614 Guided DNA Methylation Mortality Predictors 2023, 2023.11.02.23297956.<br>615 17. Geyer, P.E.; Kulak, N.A.; Pichler, G.; Holdt, L.M. 17. Geyer, P.E.; Kulak, N.A.; Pichler, G.; Holdt, L.M.; Teupser, D.; Mann, M. Plasm<br>616 Profiling to Assess Human Health and Disease. *Cell Syst* 2016, 2<br>617 doi:10.1016/j.cels.2016.02.015.<br>618 18. Völlmy, F.; van den Toor
- 618 18. Völlmy, F.; van den Toorn, H.; Zenezini Chiozzi, R.; Zucchetti, O.; Papi, A.; Volta, C.A.; For Fig. 1914, Holder, P. Holder, P. Holder, P. Holder, P. Holder, P. 185–195,<br>617 doi:10.1016/j.cels.2016.02.015.<br>618 18. Völlmy, F.; van den Toorn, H.; Zenezini Chiozzi, R.; Zucchetti, O.; Papi, A.; Volta, C.A.;<br>619 Marr 617 doi:10.1016/j.cels.2016.02.015.<br>618 18. Völlmy, F.; van den Toorn, H.; Zenezini Chiozzi, R.; Zucchetti, O.; Papi, A.; Volta, C.A.;<br>619 Marracino, L.; Vieceli Dalla Sega, F.; Fortini, F.; Demichev, V.; et al. A Serum Pr 18. Völlmy, F.; van den Toorn, H.; 2<br>619 Marracino, L.; Vieceli Dalla Sega<br>620 Signature to Predict Mortality i<br>621 e202101099, doi:10.26508/lsa.2<br>622 19. Boomsma, D.; Busjahn, A.; Pelt Marracino, L.; Vieceli Dalla Sega, F.; Fortini, F.; Demichev, V.; et al. A Serum Proteome<br>620 Signature to Predict Mortality in Severe COVID-19 Patients. *Life Sci Alliance* 2021, 4,<br>621 e202101099, doi:10.26508/Isa.202101 520 Signature to Predict Mortality in Severe COVID-19 Patients. Life Sci Alliance 2021, 4,<br>621 e202101099, doi:10.26508/lsa.202101099.<br>622 19. Boomsma, D.; Busjahn, A.; Peltonen, L. Classical Twin Studies and Beyond. Nat R
- 
- Signature to Predict Mortality in Severe COVID-19 Patients. Life Sci Alliance 2021, 4,<br>621 e202101099, doi:10.26508/lsa.202101099.<br>623 Genet 2002, 3, 872–882, doi:10.1038/nrg932.<br>624 20. Florvall, G.; Basu, S.; Larsson, A. 622 19. Boomsma, D.; Busjahn, A.; Peltonen, L. C<br>623 *Genet* 2002, 3, 872–882, doi:10.1038/nrg93<br>624 20. Florvall, G.; Basu, S.; Larsson, A. Apolipoprot<br>625 Are HDL and LDL Cholesterol for Cardiovasc<br>626 *Gerontol A Biol* 19. Boomsma, D.; Busjahn, A.; Peltonen, L. Classical Twin Studies and Beyond. Nat Rev<br>623 Genet 2002, 3, 872–882, doi:10.1038/nrg932.<br>624 20. Florvall, G.; Basu, S.; Larsson, A. Apolipoprotein A1 Is a Stronger Prognostic M
- 21. Sato, Y.; Kobayashi, T.; Nishiumi, S.; Okada, A.; Fujita, T.; Sanuki, T.; Kobayashi, M.; From The HDL and LDL Cholesterol for Cardiovascular Disease and Mortality in Elderly Men. J<br>626 627 21. Sato, Y.; Kobayashi, T.; Nishiumi, S.; Okada, A.; Fujita, T.; Sanuki, T.; Kobayashi, M.;<br>628 Asahara, M.; Adachi, M.; Gerontol A Biol Sci Med Sci 2006, 61, 1262–1266, doi:10.1093/gerona/61.12.1262.<br>
627 21. Sato, Y.; Kobayashi, T.; Nishiumi, S.; Okada, A.; Fujita, T.; Sanuki, T.; Kobayashi, M.;<br>
628 Asahara, M.; Adachi, M.; Sakai, A.; et 626 Gerontol A Biol Sci Med Sci 2006, 61, 1262–1266, doi:10.1093/gerona/61.12.1262. Asahara, M.; Adachi, M.; Sakai, A.; et al. Prospective Study Using Plasma Apolipoprotein<br>629 A2-Isoforms to Screen for High-Risk Status of Pancreatic Cancer. Cancers (Basel) 2020,<br>630 12, 2625, doi:10.3390/cancers12092625.
- 629 A2-Isoforms to Screen for High-Risk Status of Pancreatic Cancer. Cancers (Basel) 2020,<br>630 12, 2625, doi:10.3390/cancers12092625.<br>631 22. Bucki, R.; Levental, I.; Kulakowska, A.; Janmey, P.A. Plasma Gelsolin: Function, 42-Isoforms to Screen for High-Risk Status of Fancreatic Cancer. Cancers (Basel) 2020,<br>
12, 2625, doi:10.3390/cancers12092625.<br>
22. Bucki, R.; Levental, I.; Kulakowska, A.; Janmey, P.A. Plasma Gelsolin: Function,<br>
Prognost 12, 2023, doi:10.33390/cancers12092023.<br>
631 22. Bucki, R.; Levental, I.; Kulakowska, A<br>
632 Prognostic Value, and Potential Therape<br>
633 551, doi:10.2174/138920308786733912.<br>
634 23. Li, G.H.; Arora, P.D.; Chen, Y.; McCul Frognostic Value, and Potential Therapeutic Use. Curr Protein Pept Sci 2008, 9, 541–<br>633 551, doi:10.2174/138920308786733912.<br>634 23. Li, G.H.; Arora, P.D.; Chen, Y.; McCulloch, C.A.; Liu, P. Multifunctional Roles of Gelso
- 
- Frognostic Value, and Fotential Therapeutic Ose. Curr Protein Pept Scr 2000, 9, 341–<br>633 551, doi:10.2174/138920308786733912.<br>634 23. Li, G.H.; Arora, P.D.; Chen, Y.; McCulloch, C.A.; Liu, P. Multifunctional Roles of Gelso 23. Li, G.H.; Arora, P.D.; Chen, Y.; McCulloch, C635<br>
Health and Diseases. *Med Res Rev* 2012, 3<br>
636 24. McDonald, K.; Glezeva, N.; Collier, P.; O'R<br>
637 A.; Tonry, C.; Pennington, S.; Gallaghe<br>
638 Diagnostic Biomarker o Frank 1 23. Health and Diseases. *Med Res Rev* 2012, 32, 999–1025, doi:10.1002/med.20231.<br>
636 24. McDonald, K.; Glezeva, N.; Collier, P.; O'Reilly, J.; O'Connell, E.; Tea, I.; Russell-Hallinan,<br>
637 A.; Tonry, C.; Penning Freatmenta Diseases. Med Res Rev 2012, 32, 333–1023, doi:10.1002/med.20231.<br>636 24. McDonald, K.; Glezeva, N.; Collier, P.; O'Reilly, J.; O'Connell, E.; Tea, I.; Russell-Ha<br>637 A.; Tonry, C.; Pennington, S.; Gallagher, J.; A.; Tonry, C.; Pennington, S.; Gallagher, J.; et al. Tetranectin, a Potential Novel<br>638 Diagnostic Biomarker of Heart Failure, Is Expressed within the Myocardium and<br>639 Associates with Cardiac Fibrosis. *Sci Rep* 2020, 10 Fig. 2014. The energy of the state of the energy of th 638 Diagnostic Biomarker of Heart Failure, Is Expressed within the Myocardium and<br>639 Associates with Cardiac Fibrosis. Sci Rep 2020, 10, 7507, doi:10.1038/s41598-020-<br>640 64558-4.  $640$  64558-4.<br>640 64558-4.

R.; Hastie, T.; Alpert, A.; Cui, L.; et al. An Inflammatory Aging Clock (iAge) Based on<br>643 Deep Learning Tracks Multimorbidity, Immunosenescence, Frailty and Cardiovascular<br>644 Aging. Nat Aging 2021, 1, 598–615, doi:10.10

- Example 19 May 1982, The Layer Multimorbidity, Immunosenescence, Frailty and Cardiovascular<br>
643 Deep Learning Tracks Multimorbidity, Immunosenescence, Frailty and Cardiovascular<br>
645 A.M.; Engelfriet, P. Inflammatory Mark Fire Lating 2021, 1, 598–615, doi:10.1038/s43587-021-00082-y.<br>
645 26. Samson, L.D.; Buisman, A.; Ferreira, J.A.; Picavet, H.S.J.; Verschuren, W.M.M.; Boots,<br>
646 A.M.; Engelfriet, P. Inflammatory Marker Trajectories Assoc 645 26. Samson, L.D.; Buisman, A.; Ferreira, J.A.; Picavet, H.S.J.; Verschuren,<br>646 A.M.; Engelfriet, P. Inflammatory Marker Trajectories Associated with F<br>647 in a 20-year Longitudinal Study. *Clin Transl Immunology* 20<br>6 646 A.M.; Engelfriet, P. Inflammatory Marker Trajectories Associated with Frailty and Ageing<br>647 in a 20-year Longitudinal Study. *Clin Transl Immunology* 2022, 11, e1374,<br>648 doi:10.1002/cti2.1374.<br>649 27. Franceschi, C.;
- 647 in a 20-year Longitudinal Study. *Clin Transl Immunology* 2022, 11, e1374,<br>648 doi:10.1002/cti2.1374.<br>649 27. Franceschi, C.; Bonafè, M.; Valensin, S.; Olivieri, F.; De Luca, M.; Ottaviani, E.; De<br>650 Benedictis, G. I 648 doi:10.1002/cti2.1374.<br>648 doi:10.1002/cti2.1374.<br>649 27. Franceschi, C.; Bonafè, M.; Valensin, S.; Olivieri, F.; De Luca, M.; Ottaviani, E.; De<br>650 Benedictis, G. Inflamm-Aging. An Evolutionary Perspective on Immunose 649 27. Franceschi, C.; Bonafè,<br>650 Benedictis, G. Inflamm-<br>651 N Y Acad Sci 2000, 908,<br>652 28. Wischhusen, J.; Melero<br>653 From Biomarker to Nov
- 27. Franceschi, C.; Bonafè, M.; Valensin, S.; Olivieri, F.; De Luca, M.; Ottaviani, E.; De<br>
650 Benedictis, G. Inflamm-Aging. An Evolutionary Perspective on Immunosenescence. Ann<br>
651 N Y Acad Sci 2000, 908, 244–254, doi:1 650 Benedictis, G. Imlamm-Aging. An Evolutionary Perspective on Immunosenescence. Ann<br>651 N.Y.Acad Sci 2000, 908, 244–254, doi:10.1111/j.1749-6632.2000.tb06651.x.<br>652 28. Wischhusen, J.; Melero, I.; Fridman, W.H. Growth/Di 652 28. Wischhusen, J.; Melero, I.; Fridman, W.H. Growth/Differentiation Factor-1<br>653 From Biomarker to Novel Targetable Immune Checkpoint. *Front Immunol* 20<br>654 doi:10.3389/fimmu.2020.00951.<br>655 29. Liu, H.; Huang, Y.; L From Biomarker to Novel Targetable Immune Checkpoint. Front Immunol 2020, 11, 951,<br>654 doi:10.3389/fimmu.2020.00951.<br>655 29. Liu, H.; Huang, Y.; Lyu, Y.; Dai, W.; Tong, Y.; Li, Y. GDF15 as a Biomarker of Ageing.<br>656 Experi
- From Biomarker to Novel Targetable Immune Checkpoint. Front Immunol 2020, 11, 991,<br>654 doi:10.3389/fimmu.2020.00951.<br>655 29. Liu, H.; Huang, Y.; Lyu, Y.; Dai, W.; Tong, Y.; Li, Y. GDF15 as a Biomarker of Ageing.<br>656 *Exper* 655 29. Liu, H.; Huang, Y.; Lyu, Y.; Dai, Y.; Experimental Gerontology 2021, .<br>656 Experimental Gerontology 2021, .<br>657 30. Wan, Y.; Fu, J. GDF15 as a Key<br>658 Diseases and Ageing. *Mol Cell Bio*<br>659 31. de Gonzalo-Calvo, D
- 
- Experimental Gerontology 2021, 146, 111228, doi:10.1016/j.exger.2021.111228.<br>657 30. Wan, Y.; Fu, J. GDF15 as a Key Disease Target and Biomarker: Linking Chronic Lung<br>658 Diseases and Ageing. Mol Cell Biochem 2023, 1–14, d Experimental Gerontology 2021, 146, 111228, doi:10.1010/j.exger.2021.111228.<br>657 30. Wan, Y.; Fu, J. GDF15 as a Key Disease Target and Biomarker: Linking Chroni<br>658 Diseases and Ageing. *Mol Cell Biochem* 2023, 1–14, doi:1 Example 2013, 1-14, doi:10.1007/s11010-023-04743-x.<br>
658 Diseases and Ageing. *Mol Cell Biochem* 2023, 1-14, doi:10.1007/s11010-023-04743-x.<br>
659 31. de Gonzalo-Calvo, D.; Neitzert, K.; Fernández, M.; Vega-Naredo, I.; Caba Experience Molecule Material Ageing. Molecule Material Color, 1944, 00.110.1007, 311010-023-04743-x.<br>
659 31. de Gonzalo-Calvo, D.; Neitzert, K.; Fernández, M.; Vega-Naredo, I.; Caballero, B.; García<br>
660 Macía, M.; Suárez For Columbus Columbus Columbus Columbia and Disease: TNF-Alpha and IL-6 as<br>
For Possible Biomarkers. Free Radic Biol Med 2010, 49, 733–737,<br>
doi:10.1016/j.freeradbiomed.2010.05.019.<br>
S64 32. Maggio, M.; Guralnik, J.M.; Lon Free Radic Biol Med 2010, 19, 733–737, doi:10.1016/j.freeradbiomed.2010.05.019.<br>663 doi:10.1016/j.freeradbiomed.2010.05.019.<br>664 32. Maggio, M.; Guralnik, J.M.; Longo, D.L.; Ferrucci, L. Interleukin-6 in Aging and Chronic<br>
- 664 32. Maggio, M.; Guralnik, J.M.; Longo, D.L.; Ferrucci, L. Interleukin-6 in Aging and Chronic
- Fossible Biomarkers. Thee Radic Biol Med 2010, 49, 733–737,<br>663 doi:10.1016/j.freeradbiomed.2010.05.019.<br>664 32. Maggio, M.; Guralnik, J.M.; Longo, D.L.; Ferrucci, L. Interleukin-6 in Aging and Chronic<br>665 Disease: A Magni 664 32. Maggio, M.; Guralnik, J.M.; Longo, D.L.; Fe<br>665 Disease: A Magnificent Pathway. *J Gerontol*<br>666 33. Valiathan, R.; Ashman, M.; Asthana, D. Effec<br>667 to Elderly. *Scand J Immunol* **2016**, 83, 255–2<br>668 34. Tavares,
- Example 18 Disease: A Magnificent Pathway. J Gerontol A Biol Sci Med Sci 2006, 61, 575–584.<br>
666 33. Valiathan, R.; Ashman, M.; Asthana, D. Effects of Ageing on the Immune System: Infants<br>
to Elderly. *Scand J Immunol* 201 666 Bisease: A Magnificent Pathway. J Gerontol A Biol Scribted Scr 2000, 01, 373–384.<br>666 Biol Sci Mediathan, R.; Ashman, M.; Asthana, D. Effects of Ageing on the Immune System: I<br>667 to Elderly. *Scand J Immunol* 2016, 83 to Elderly. *Scand J Immunol* **2016**, 83, 255–266, doi:10.1111/sji.12413.<br>668 34. Tavares, S.M.Q.M.C.; Junior, W. de L.B.; Lopes e Silva, M.R. Normal Lymphocyte<br>669 Immunophenotype in an Elderly Population. *Rev Bras Hemat* 668 34. Tavares, S.M.Q.M.C.; Junior, W. de L.B.; Lopes e Silva, M.R. Nor<br>669 Immunophenotype in an Elderly Population. *Rev Bras Hematol Hemote*<br>670 183, doi:10.1016/j.bjhh.2014.03.021.<br>671 35. Yan, J.; Greer, J.M.; Hull,
- 1069 11 Immunophenotype in an Elderly Population. *Rev Bras Hematol Hemoter* 2014, 36, 180–<br>183, doi:10.1016/j.bjhh.2014.03.021.<br>671 35. Yan, J.; Greer, J.M.; Hull, R.; O'Sullivan, J.D.; Henderson, R.D.; Read, S.J.; McComb France II and Elderly Population. *Nev Bras Hematol Hemoter 2014*, 36, 180–183, doi:10.1016/j.bjhh.2014.03.021.<br>671 35. Yan, J.; Greer, J.M.; Hull, R.; O'Sullivan, J.D.; Henderson, R.D.; Read, S.J.; McCombe, P.A.<br>672 The E 671 35. Yan, J.; Greer, J.M.; Hull, R.; O'Sullivan<br>672 The Effect of Ageing on Human Ly<br>673 Females. *Immunity & Ageing* 2010, 7,<br>674 36. Nikolich-Žugich, J. Ageing and Life-Lo<br>675 Latent Persistent Infections. *Nat Rev In* 185. Yan, J.; Greer, J.M.; Hull, R.; O'Sullivan, J.D.; Henderson, R.D.; Read, S.J.; McCombe, P.A.<br>
1672 The Effect of Ageing on Human Lymphocyte Subsets: Comparison of Males and<br>
1673 Females. *Immunity & Ageing* **2010**, 7
- 
- Females. *Immunity & Ageing* 2010, 7, 4, doi:10.1186/1742-4933-7-4.<br>674 36. Nikolich-Žugich, J. Ageing and Life-Long Maintenance of T-Cell Subsets in the Face of<br>675 Latent Persistent Infections. *Nat Rev Immunol* 2008, 8, 673 Females. *Immunity* & Ageing 2010, 7, 4, doi:10.1180/1742-4933-7-4.<br>674 36. Nikolich-Žugich, J. Ageing and Life-Long Maintenance of T-Cell Subs<br>675 Latent Persistent Infections. Nat Rev Immunol 2008, 8, 512–522, doi:1<br> Latent Persistent Infections. Nat Rev Immunol 2008, 8, 512–522, doi:10.1038/nri2318.<br>676 37. Schoenmaker, M.; de Craen, A.J.M.; de Meijer, P.H.E.M.; Beekman, M.; Blauw, G.J.;<br>677 Slagboom, P.E.; Westendorp, R.G.J. Evidence Eatent Persistent Infections: *Nut Nev Immunol* 2008, 8, 312–322, doi:10.1038/mi2318.<br>676 37. Schoenmaker, M.; de Craen, A.J.M.; de Meijer, P.H.E.M.; Beekman, M.; Blauw, G.J<br>678 Survival Using a Family Approach: The Leiden Slagboom, P.E.; Westendorp, R.G.J. Evidence of Genetic Enrichment for Exceptional<br>678 Survival Using a Family Approach: The Leiden Longevity Study. *European Journal of*<br>679 Human Genetics 2006, 14, 79–84, doi:10.1038/sj.e
- Survival Using a Family Approach: The Leiden Longevity Study. *European Journal of*<br>679 Human Genetics 2006, 14, 79–84, doi:10.1038/sj.ejhg.5201508.<br>680 38. Ikram, M.A.; Brusselle, G.G.O.; Murad, S.D.; van Duijn, C.M.; Fra 679 Survival Using a Family Approach: The Leiden Longevity Study. European Journal of<br>680 38. Ikram, M.A.; Brusselle, G.G.O.; Murad, S.D.; van Duijn, C.M.; Franco, O.H.;<br>681 Goedegebure, A.; Klaver, C.C.W.; Nijsten, T.E.C. Framan Genetics 2000, 14, 79–84, doi:10.1038/sj.ejhg.5201508.<br>680 38. Ikram, M.A.; Brusselle, G.G.O.; Murad, S.D.; van Duijn,<br>681 Goedegebure, A.; Klaver, C.C.W.; Nijsten, T.E.C.; Peeters, R.P.; S<br>682 Rotterdam Study: 2018 681 Goedegebure, A.; Klaver, C.C.W.; Nijsten, T.E.C.; Peeters, R.P.; Stricker, B.H.; et al. The<br>682 Rotterdam Study: 2018 Update on Objectives, Design and Main Results. *Eur J Epidemiol*<br>**2017**, 32, 807–850, doi:10.1007/s1
- Rotterdam Study: 2018 Update on Objectives, Design and Main Results. *Eur J Epidemiol*<br>683 **2017**, 32, 807–850, doi:10.1007/s10654-017-0321-4.<br>684 39. Würtz, P.; Kangas, A.J.; Soininen, P.; Lawlor, D.A.; Davey Smith, G.; A For the Rotterdam Study: 2018 Optate on Objectives, Design and Main Results. Eur J Epidemiol 683<br> **2017**, 32, 807–850, doi:10.1007/s10654-017-0321-4.<br>
684 39. Würtz, P.; Kangas, A.J.; Soininen, P.; Lawlor, D.A.; Davey Smit **2017, 32, 807–850, doi:10:1007,310054-017-0321-4.**<br>684 39. Würtz, P.; Kangas, A.J.; Soininen, P.; Lawlor, D.A.; I<br>685 Quantitative Serum Nuclear Magnetic Resonanc<br>686 Epidemiology: A Primer on -Omic Technologies. An<br>687 1 Constitution of the United Serum Nuclear Magnetic Resonance Metabolomics in Large-Scale<br>686 5 Epidemiology: A Primer on -Omic Technologies. Am. J. Epidemiol. 2017, 186, 1084–<br>1096, doi:10.1093/aje/kwx016. Epidemiology: A Primer on -Omic Technologies. Am. J. Epidemiol. 2017, 186, 1084–1096, doi:10.1093/aje/kwx016.  $687$  Epidemiology: A Primer on -Omic Technologies. Am. J. Epidemiol. 2017, 186, 1884–1096, doi:10.1093/aje/kwx016. 687 1096, doi:10.1093/aje/kwx016.

Muclear Magnetic Resonance Metabolomics in Cardiovascular Epidemiology and<br>690 6enetics. Circ Cardiovasc Genet 2015, 8, 192–206,<br>691 doi:10.1161/CIRCGENETICS.114.000216.<br>692 41. Scholman, R.C.; Giovannone, B.; Hiddingh, S.

- 680 Genetics. *Circ* Cardiovasc Genet 2015, 8, 192–206,<br>691 doi:10.1161/CIRCGENETICS.114.000216.<br>692 41. Scholman, R.C.; Giovannone, B.; Hiddingh, S.; Meerding, J.M.; Malvar Fernandez, B.;<br>693 van Dijk, M.E.A.; Tempelman, Genetics. Circ Cardiovasc Genet 2015, 0, 132–200,<br>691 doi:10.1161/CIRCGENETICS.114.000216.<br>692 41. Scholman, R.C.; Giovannone, B.; Hiddingh, S.; Meerding, J.M.; Malvar Fernandez, B.;<br>693 van Dijk, M.E.A.; Tempelman, M.J.; 41. Scholman, R.C.; Giovannone, B.; Hiddin,<br>693 van Dijk, M.E.A.; Tempelman, M.J.; Prak<br>694 on 162 Circulating Immune Related Prot<br>695 114–124, doi:10.1016/j.cyto.2017.10.021<br>696 42. Skowronek, P.; Thielert, M.; Voytik, E. For March Hall, M.E.A.; Tempelman, M.J.; Prakken, B.J.; de Jager, W. Effect of Anticoagulants<br>
694 on 162 Circulating Immune Related Proteins in Healthy Subjects. Cytokine 2018, 106,<br>
114–124, doi:10.1016/j.cyto.2017.10.02
- on 162 Circulating Immune Related Proteins in Healthy Subjects. Cytokine 2018, 106,<br>695 114–124, doi:10.1016/j.cyto.2017.10.021.<br>696 42. Skowronek, P.; Thielert, M.; Voytik, E.; Tanzer, M.C.; Hansen, F.M.; Willems, S.; Kar 695 on 162 chediating immune Related Proteins in Healthy Subjects. Cytokine 2018, 100,<br>696 42. Skowronek, P.; Thielert, M.; Voytik, E.; Tanzer, M.C.; Hansen, F.M.; Willems, S.; Karayel,<br>697 0.; Brunner, A.-D.; Meier, F.; M 42. Skowronek, P.; Thielert, M.; Voytik, E.; Tan<br>697 0.; Brunner, A.-D.; Meier, F.; Mann, M. Ra<br>698 Proteome With Deep Libraries and Optii<br>699 Proteomics 2022, 21, 100279, doi:10.1016,<br>700 43. Demichev, V.; Messner, C.B.; 697 0.; Brunner, A.-D.; Meier, F.; Mann, M. Rapid and In-Depth Coverage of the (Phospho-<br>698 Proteome With Deep Libraries and Optimal Window Design for Dia-PASEF. *Mol Cell*<br>699 Proteomics 2022, 21, 100279, doi:10.1016/j.m
- 702 Throughput. Nat Methods 2020, 17, 41-44, doi:10.1038/s41592-019-0638-x. Froteomics 2022, 21, 100279, doi:10.1016/j.mcpro.2022.100279.<br>
1699 Proteomics 2022, 21, 100279, doi:10.1016/j.mcpro.2022.100279.<br>
1699 Proteomics 2022, 21, 100279, doi:10.1016/j.mcpro.2022.100279.<br>
1699 Proteomics 2022, 2 Froteomics 2022, 21, 100279, doi:10.10107].incpio:2022.100279.<br>
700 43. Demichev, V.; Messner, C.B.; Vernardis, S.I.; Lilley, K.S.; Ralser<br>
701 Networks and Interference Correction Enable Deep Proteom<br>
702 Throughput. Nat
- To 1 Networks and Interference Correction Enable Deep Proteome Coverage in High<br>
702 Throughput. Nat Methods 2020, 17, 41–44, doi:10.1038/s41592-019-0638-x.<br>
703 44. Bizzarri, D.; Reinders, M.J.T.; Beekman, M.; Slagboom, P Throughput. Nat Methods 2020, 17, 41–44, doi:10.1038/s41592-019-0638-x.<br>
703 44. Bizzarri, D.; Reinders, M.J.T.; Beekman, M.; Slagboom, P.E.; van den Akker, E.B. MiMIR:<br>
704 R-Shiny Application to Infer Risk Factors and En 11100ghput. Nat Methods 2020, 17, 41–44, doi:10.1038/s41552-015-0058-x.<br>
703 44. Bizzarri, D.; Reinders, M.J.T.; Beekman, M.; Slagboom, P.E.; van den Akker, E.<br>
704 R-Shiny Application to Infer Risk Factors and Endpoints f 44. Bizzarri, D.; Reinders, M.J.T.; Beekman, M.; Slagboom, P.E.; van den Akker, E.B. MiMIR:<br>
704 R-Shiny Application to Infer Risk Factors and Endpoints from Nightingale Health's 1H-<br>
705 MMR Metabolomics Data. *Bioinforma*
- Bioinformatics 2013, 14, 128, doi:10.1186/1471-2105-14-128. 105 NMR Metabolomics Data. Biomformatics 2022, 38, 3847–3849,<br>
106 doi:10.1093/bioinformatics/btac388.<br>
107 45. Chen, E.Y.; Tan, C.M.; Kou, Y.; Duan, Q.; Wang, Z.; Meirelles, G.V.; Clark, N.R.; Ma'ayan, A.<br>
108 Enrichr: In 107 45. Chen, E.Y.; Tan, C.M.; Kou, Y.; Duan, Q.<br>108 Enrichr: Interactive and Collaborative<br>109 Bioinformatics **2013**, 14, 128, doi:10.1<br>10 46. Sayols, S. Rrvgo: A Bioconductor Pack<br>11 MicroPubl Biol 20
- For the chem, E.Y., E.W., E.Y., E.W., E.Y., E.W., Chen, E.W., Chen, Chen, Chen, Chen, Chen, Chen, Chen, Chen, C<br>
709 Bioinformatics 2013, 14, 128, doi:10.1186/1471-2105-14-128.<br>
710 A6. Sayols, S. Rrvgo: A Bioconductor Pac Finch: Interactive and Collaborative HTML5 Gene List Enrichment Analysis 1001. BMC<br>
709 Bioinformatics 2013, 14, 128, doi:10.1186/1471-2105-14-128.<br>
710 46. Sayols, S. Rrvgo: A Bioconductor Package for Interpreting Lists o Prompomatics 2013, 14, 128, doi:10.1186/1471-2105-14-128.<br>
710 46. Sayols, S. Rrvgo: A Bioconductor Package for Interpreting Lists<br>
711 MicroPubl Biol 2023, 10.17912/r<br>
712 doi:10.17912/micropub.biology.000811.<br>
713 711 MicroPubl Biol 2023, 10.17912/micropub.biology.000811,<br>712 doi:10.17912/micropub.biology.000811.<br>713<br>715  $\frac{600}{2023}$ , 10.17912/micropub.biology.000811.<br>
712 doi:10.17912/micropub.biology.000811.<br>
713<br>
715 713<br>713<br>714<br>715
- 
- 713
- 715

716 Figures Legends<br>717 Figure 1: Study descripti<br>718 each study population. B<br>719 within each cohort.<br>720 Figure 2: Increase GDF15 Figure 1: Study description. A) Flowchart detailing the inclusion criteria for samples within<br>
218 each study population. B) Contrast in MetaboHealth, lymphocyte percentage, age and BMI<br>
720 **Figure 2: Increase GDF15, IL6,** 

within each cohort.<br>720<br>721 **Figure 2: Increase GDF15, IL6, and MIG associate with MetaboHealth levels.** A) Differential<br>722 detectability between cases (top) and controls (bottom) for the Luminex cytokine measures,<br>723 wi The Matter Controller<br>721 **Figure 2: Increase G**<br>722 detectability betwee<br>723 with detected value<br>724 adjusted p-value 720<br>721<br>722<br>723<br>724<br>725 detectability between cases (top) and controls (bottom) for the Luminex cytokine measures,<br>
723 with detected values in blue and undetected in grey. The heatmap on the bottom details the<br>
724 adjusted p-value of the Fisher with detected values in blue and undetected in grey. The heatmap on the bottom details the<br>
224 adjusted p-value of the Fisher test evaluating the significances of the differential<br>
225 detectability. C) Volcano-plot perta 224 adjusted p-value of the Fisher test evaluating the significances of the differential<br>725 detectability. C) Volcano-plot pertaining the univariate associations between the cytokines'<br>726 levels and the participants' cas detectability. C) Volcano-plot pertaining the univariate associations between the cytokines'<br>726 levels and the participants' case/control status corrected for sex, and age. In red the<br>727 positively associated cytokines, The Samuel Case of the participants' case/control status corrected for sex, and age. In red the<br>
The positively associated cytokines, in grey the ones not significant. D) Heatmap depicting the<br>
The correlations between eac Figure 227 and the participants and the participants and the participant. D) Heatmap depicting the<br>
728 correlations between each cytokines with the components of the MetaboHealth score.<br>
730 **Figure 3: Quantitative plasma** 

Correlations between each cytokines with the components of the MetaboHealth score.<br>
729<br> **Figure 3: Quantitative plasma proteomics reveals proteins that associate with<br>
731 MetaboHealth scores in RS and LLS-PAROFFS cohorts** 729<br>730 Figure 3: Quantitative plasma proteomics reveals proteins that associate<br>731 MetaboHealth scores in RS and LLS-PAROFFS cohorts. A) Volcano plot depicting associa<br>732 of proteins with MetaboHealth scores (corrected 729<br>730<br>731<br>732<br>733<br>734<br>735 The Tigure 3: Quantitative plasma proteomics reveals proteins that associate with<br>731 MetaboHealth scores in RS and LLS-PAROFFS cohorts. A) Volcano plot depicting associations<br>732 of proteins with MetaboHealth scores (corr The MetaboHealth scores in NS and LLS-PAROFFS collors. A) volcano plot depicting associations<br>1732 of proteins with MetaboHealth scores (corrected for sex, and age). In blue the negatively<br>1733 and in red the positively as and in red the positively associated plasma proteins are depicted. Grouped significantly<br>
T34 enriched Gene Ontology Biological Processes are shown for B) the negatively and C)<br>
T35 positively associated plasma proteins. D The intiched Gene Ontology Biological Processes are shown for B) the negatively and C)<br>
The positively associated plasma proteins. D) and E) depict the enriched KEGG pathways for the<br>
The plasma proteins respectively negat The Contrology Biology Biology<br>
735 positively associated plasma proteins. D) and E) depict the enriched KEGG pathways for the<br>
736 plasma proteins respectively negatively and positively associated with the MetaboHealth<br>
7 positively associated with the MetaboHealth<br>137 score.<br>138 **Figure 4: Plasma proteins associated with extreme MetaboHealth scores validated in the<br>140 NTR dataset. A) Profile depicting the values of the most significantly** 

Figure 2. Thas provides a section of the setting of the STAR State of the STAR STAR Profile depicting the values of the most significantly associated proteins<br>1986 **Profile 1989 Profile STAR Profile depicting the Values of** 738<br>739 **Figure**<br>740 **NTR d**<br>741 (|estin<br>742 contro 739<br>740<br>741<br>742<br>743<br>744 Tigure 4: Plasma proteins associated with extreme Metaborieanth scores vanitated in the<br>
740 **NTR dataset.** A) Profile depicting the values of the most significantly associated proteins<br>
741 (|estimate|>1) and cytokines (y 740 **NTR dataset.** A) Frome depicting the values of the most significantly associated proteins<br>741 (|estimate|>1) and cytokines (y-axis) in all samples (x-axis), clearly separating the cases and<br>742 controls in all 3 datas Frammerical and the stimate of the stimate of the show the phenotypic characteristics of<br>143 all individuals. The annotation on the left shows the associations' estimate, and log10(FDR)<br>144 observed in RS and LLS, and the all individuals. The annotation on the left shows the associations' estimate, and log10(FDR)<br>
744 observed in RS and LLS, and the platform for each feature (Mass Spectrometry based Plasma<br>
745 Proteomics or Luminex). B) Be The all individuals. The annotation of the platform for each feature (Mass Spectrometry based Plasma<br>
245 Proteomics or Luminex). B) Beta-beta plot comparing the linear models in the extremes of<br>
246 RS and LLS\_PAROFFs on Proteomics or Luminex). B) Beta-beta plot comparing the linear models in the extremes of<br>T46 RS and LLS\_PAROFFs on the x-axis and the linear mixed models in the NTR on the y-axis. Each<br>dots corresponds to a feature, and it 746 RS and LLS\_PAROFFs on the x-axis and the linear mixed models in the NTR on the y-axis. Each<br>747 dots corresponds to a feature, and it's colored based on significance. The labels are shown<br>748 only when the features are dots corresponds to a feature, and it's colored based on significance. The labels are shown 2748 only when the features are consistently significant in the 2 analyses.<br>
749<br>
750<br>
751 **Supplementary Figures legends** 

# 749<br>750<br>751 **Supplementary Figures legends**<br>752 **only Figure S1: Ouality control of the cytokines in LLS-PAROFFS**

750<br>751<br>752<br>753<br>754 751<br>752<br>753<br>754<br>755 **Supplementary Figures legends**<br>
752 **Figure S1: Quality control of the cytokines in L**<br>
754 of undetected values for each cytokine (x-axis) div<br>
755 detectability analysis performed separately for B) I<br>
756 distributions 753<br>754<br>755<br>757<br>757<br>758 Figure S1: Quality control of the cytokines in LLS-PAROFFS and RS. A) Percentage<br>of undetected values for each cytokine (x-axis) divided per cohort (y-axis). Differential<br>detectability analysis performed separately for B) 2754 of undetected values for each cytokine (x-axis) divided per cohort (y-axis). Differential<br>
2755 detectability analysis performed separately for B) LLS-PAROFFS and C) RS. Cytokines<br>
2756 distributions C) before and D)

detectability analysis performed separately for B) LLS-PAROFFS and C) RS. Cytokines<br>
756 distributions C) before and D) after the pre-processing steps (log transform and z-scaling).<br>
757 **Figure S2: Sensitivity analyses an** distributions C) before and D) after the pre-processing steps (log transform and z-scaling).<br>
757<br> **Figure S2: Sensitivity analyses and Meta-analyses of the cytokines' associations w<br>
MetaboHealth extremes:** A- D) Volcano-759<br>760<br>761 **Figure S2: Sensitivity analyses and Meta-analyses of the cytokines' associations with**<br>T59 **MetaboHealth extremes:** A- D) Volcano-plot of the univariate linear associations between<br>the cytokines and the case/control statu 759 **MetaboHealth extremes:** A- D) Volcano-plot of the univariate linear associations between<br>760 the cytokines and the case/control status of the participants systematically corrected by<br>761 increasing covariates. In grey 760 the cytokines and the case/control status of the participants systematically corrected by increasing covariates. In grey the not significant cytokines and red the positively significant increasing covariates. In grey t 761 increasing covariates. In grey the not significant cytokines and red the positively significant

The metalliers ones. E) Intercorrelations of all the cytokines. F) Forest plots of the univariate linear<br>
763 associations divided in RS (green), LLS-PAROFFS (yellow), and Meta-analyses (red).<br>
765 **Figure S3: Quality cont** associations divided in RS (green), LLS-PAROFFS (yellow), and Meta-analyses (red).<br>
764<br> **Figure S3: Quality control of the proteomics dataset in RS and LLS-PAROF**<br>
766 Percentage of missingness in the highly missing featu 764<br>765<br>766<br>767<br>768<br>769<br>770 Figure S3: Quality control of the proteomics dataset in RS and LLS-PAROFFS: A)<br>
766 Percentage of missingness in the highly missing features divided per cohort. B) Differential<br>
767 detectability between cases (top) and co Percentage of missingness in the highly missing features divided per cohort. B) Differential<br>
767 detectability between cases (top) and controls (bottom) for the proteins showing significant<br>
768 difference. The heatmap on difference. The heatmap on the bottom details the adjusted p-value of the Fisher test<br>
769 evaluating the significance of the differential detectability. 9 randomly selected plasma<br>
770 proteins distributions C) before and evaluating the significance of the differential detectability. 9 randomly selected plasma<br>
770 proteins distributions C) before and D) after the pre-processing steps (log transform and z-<br>
771 saling).<br>
772 **Figure S4: Sen** 

proteins distributions C) before and D) after the pre-processing steps (log transform and z-<br>
771 scaling).<br> **Tigure S4: Sensitivity analyses and Meta-analyses of the plasma proteins with the<br>
MetaboHealth extremes in RS a** 771 scaling).<br>
772 **Figure<br>
773 <b>Figure<br>
775** linear as<br>
776 systemati 772<br>773<br>774<br>775<br>776<br>777<br>778 **Figure S4: Sensitivity analyses and Meta-analyses of the plasma proteins with the**<br> **MetaboHealth extremes in RS and LLS-PAROFFS:** A- D) Volcano-plot of the univariate<br>
linear associations between the cytokines and the ca **MetaboHealth extremes in RS and LLS-PAROFFS:** A- D) Volcano-plot of the univariate<br>
linear associations between the cytokines and the case/control status of the participants<br>
systematically corrected by increasing covaria 776 systematically corrected by increasing covariates. In grey the not significant cytokines, in<br>
777 blue the negatively and red the positively significant plasma proteins. Upset plot detailing the<br>
778 amount of signific systematically corrected by increasing covariates. In grey the not significant cytokines, in<br>
777 blue the negatively and red the positively significant plasma proteins. Upset plot detailing the<br>
778 amount of significantl blue the negatively and red the positively significant plasma proteins. Upset plot detailing the<br>
1778 amount of significantly D) negatively and E) positively associated features after correcting<br>
179 for covariates.<br>
179

The amount of significantly D) negatively and E) positively associated features after correcting<br>
The form covariates.<br>
The GS: Meta-analyses of the associations between the proteomics features and the<br>
The MetaboHealth ex 779 for covariates.<br>
780<br>
781 **Figure S5:**<br>
782 **MetaboHealth**<br>
783 and B) positive<br>
784<br>
795 **Figure S6:** 782<br>783<br>784<br>785<br>786 **Figure S5: Meta-analyses of the associations between the proteomics features and the**<br> **TREP MetaboHealth extremes:** Forest-plot of the meta-analysis for the significantly A) Negatively<br> **TREP 34**<br> **Figure S6: Correlation** 

**MetaboHealth extremes:** Forest-plot of the meta-analysis for the significantly A) Negatively<br>
783 and B) positively associated features.<br> **Figure S6: Correlations of the significantly associated proteomics:** Heatmaps of t 783 and B) positively associated features.<br>784<br>785 **Figure S6: Correlations of the s**<br>786 correlations of the proteomics feature<br>787 with the proteomics features themsely<br>788 **Figure S6: Correlations of the significantly associated proteomics:** Heatmaps of the correlations of the proteomics features with A) MetaboHealth and its components, and B) with the proteomics features themselves.<br> **Figur** 

787<br>788<br>789<br>790 correlations of the proteomics features with A) MetaboHealth and its components, and B)<br>
787 with the proteomics features themselves.<br>
788 **Figure 57: Functional Enrichments of the proteomics datasets:** Bar-plots displayin 787 with the proteomics features themselves.<br>
789 Figure S7: Functional Enrichments of<br>
790 log10pvalues of the significant enrichme<br>
791 A) GO Biological Processes pathways,<br>
792 significantly enriched in the C) negativel 791<br>792<br>793<br>794 Figure S7: Functional Enfidences of the procedules datasets: Bar-plots displaying the<br>1990 Figure S7: Functional Enrichments performed for all 320 features in the platform in<br>191 A) GO Biological Processes pathways, and B) 200 Biological Processes pathways, and B) KEGG pathways. GO Biological Processes<br>792 significantly enriched in the C) negatively and D) positively associated proteomics features.<br>793 Reactome pathways significantly enriche The significantly enriched in the C) negatively and D) positively associated proteomics features.<br>
The Reactome pathways significantly enriched in the E) negatively and F) positively associated<br>
The Proteomics features.<br>
T The Reactome pathways significantly enriched in the E) negatively and F) positively associated<br>
The Proteomics features.<br>
The Tigure S8: Covid related proteins. Plots displaying the case/control (red/blue) differences<br>
The

Figure Cattor pathways significantly entities in the E) negatively and F) positively are entered<br>195<br>196 **Figure S8: Covid related proteins.** Plots displaying the case/control (red/blue) differences<br>197 in RS and LLS-PAROF 795<br>796 **Figure S8: Covid m**<br>797 **in RS and LLS-PAROFI**<br>798 al. Figure S8: Covid related proteins. Plots displaying the case/control (red/blue) differences<br>797 in RS and LLS-PAROFFs for the proteins previously associated to covid infection by Völmy et<br>798 al.<br>800 **Figure S9: Pre-proces** 

795<br>796<br>797<br>798<br>799<br>800<br>801

798 al.<br>799 **Figure S9: Pre-processing in NTR.** A) Percentage of undetected values per cytokines (y-axis) per sample (x-axis). B) Percentage of undetected values per proteomics feature (y-axis) in each sample (x-axis). C) 799<br>799<br>800<br>801 axi<br>802 in ( 802<br>803 Figure S9: Pre-processing in NTR. A) Percentage of undetected values per cytokines (y-<br>
801 axis) per sample (x-axis). B) Percentage of undetected values per proteomics feature (y-axis)<br>
802 in each sample (x-axis). C) Dif

axis) per sample (x-axis). B) Percentage of undetected values per proteomics feature (y-axis)<br>
802 in each sample (x-axis). C) Differential detectability of the cytokines in NTR.<br>
803<br> **Figure S10: Associations in NTR.** Vo 802 in each sample (x-axis). C) Differential detectability of the cytokines in NTR.<br>803<br>804 **Figure S10: Associations in NTR.** Volcano plot of the univariate linear<br>806 comparing the case/control status with A) cytokines a 804<br>805<br>806<br>807<br>808<br>809 805 rigue S10: Associations in NTR. Volcano plot of the univariate linear mixed models<br>806 comparing the case/control status with A) cytokines and B) proteomics features corrected<br>807 for age, and sex. Upset plots of the signi 807 for age, and sex. Upset plots of the significantly C) negative and D) positive proteomics markers.<br>808 markers.<br>809 ex. Upset plots of the significant sex. Upset plots of the significantly C) negative and D) positive proteinant<br>808 markers.<br>809 809<br>809

810 Figure S11: Complete Heatmap profiles of the significant proteins. Profile depicting the<br>811 values of the significantly associated proteins (FDR<0.05) and cytokines (y-axis) in all<br>812 samples (x-axis), clearly separa samples (x-axis), clearly separating the cases and controls in all 3 datasets. The annotations For the case of the case of the cases and the cases and individuals. The annotation on the left<br>
813 on the top show the phenotypic characteristics of all individuals. The annotation on the left<br>
815 for each feature (Mass Section the top shows the associations' estimate, and log10(FDR) observed in RS and LLS, and the platform<br>815 for each feature (Mass Spectrometry or Luminex).<br>816 815 Shows the associations of the associations' estimately and log10(FDR) observed in RS and LLS, and the platform<br>816<br>816

 $816$ 

Figure 1: MetaboHealth-driven selection and main phenotypic differences



control cases

## **Figure 2:** Increased GDF15, IL6 and MIG associate with higher MetaboHealth levels



**Figure 3:** 68 negatively and 38 positively associated serum proteins



