

## ARTICLE OPEN



# A polygenic risk score for multiple myeloma risk prediction

Federico Canzian<sup>1✉</sup>, Chiara Piredda<sup>1,2</sup>, Angelica Macaudo<sup>1,2</sup>, Daria Zawirska<sup>3</sup>, Niels Frost Andersen<sup>4</sup>, Arnon Nagler<sup>5</sup>, Jan Maciej Zaucha<sup>6</sup>, Grzegorz Mazur<sup>7</sup>, Charles Dumontet<sup>8</sup>, Marzena Wątek<sup>9</sup>, Krzysztof Jamrozik<sup>10</sup>, Juan Sainz<sup>11,12</sup>, Judit Várkonyi<sup>13</sup>, Aleksandra Butrym<sup>14</sup>, Katia Beider<sup>5</sup>, Niels Abildgaard<sup>15</sup>, Fabienne Lesueur<sup>16</sup>, Marek Dudziński<sup>17</sup>, Annette Juul Vangsted<sup>18</sup>, Matteo Pelosini<sup>19</sup>, Edyta Subocz<sup>20</sup>, Mario Petrini<sup>19</sup>, Gabriele Buda<sup>19</sup>, Małgorzata Rażny<sup>21</sup>, Federica Gemignani<sup>2</sup>, Herlander Marques<sup>22</sup>, Enrico Orciuolo<sup>19</sup>, Katalin Kadar<sup>13</sup>, Artur Jurczynszyn<sup>23</sup>, Agnieszka Druzd-Sitek<sup>24</sup>, Ulla Vogel<sup>25</sup>, Vibeke Andersen<sup>26</sup>, Rui Manuel Reis<sup>22,27,28</sup>, Anna Suska<sup>23</sup>, Hervé Avet-Loiseau<sup>29</sup>, Marcin Kruszewski<sup>30</sup>, Waldemar Tomczak<sup>31</sup>, Marcin Rymko<sup>32</sup>, Stephane Minvielle<sup>33</sup> and Daniele Campa<sup>2</sup>

© The Author(s) 2021

There is overwhelming epidemiologic evidence that the risk of multiple myeloma (MM) has a solid genetic background. Genome-wide association studies (GWAS) have identified 23 risk loci that contribute to the genetic susceptibility of MM, but have low individual penetrance. Combining the SNPs in a polygenic risk score (PRS) is a possible approach to improve their usefulness. Using 2361 MM cases and 1415 controls from the International Multiple Myeloma rESEarch (IMMEnSE) consortium, we computed a weighted and an unweighted PRS. We observed associations with MM risk with OR = 3.44, 95% CI 2.53–4.69,  $p = 3.55 \times 10^{-15}$  for the highest vs. lowest quintile of the weighted score, and OR = 3.18, 95% CI 2.1 = 34–4.33,  $p = 1.62 \times 10^{-13}$  for the highest vs. lowest quintile of the unweighted score. We found a convincing association of a PRS generated with 23 SNPs and risk of MM. Our work provides additional validation of previously discovered MM risk variants and of their combination into a PRS, which is a first step towards the use of genetics for risk stratification in the general population.

*European Journal of Human Genetics* (2022) 30:474–479; <https://doi.org/10.1038/s41431-021-00986-8>

## INTRODUCTION

Multiple myeloma (MM) is the third most common hematological malignancy with a worldwide incidence rate of 2.1/100,000 new cases each year (<https://gco.iarc.fr/today/home>) [1]. MM is preceded by monoclonal gammopathy of undetermined significance (MGUS), an asymptomatic premalignant condition [2, 3], and by smoldering myeloma (SM), a more advanced precursor of the disease [4].

MM etiology has a strong genetic component, with several variants associated with its risk [5–21]. In particular, genome-wide associations studies (GWAS) identified 23 MM risk loci, but as for many other traits the individual penetrance of each SNP is low, with odds ratios (OR) per risk allele ranging from 1.11 to 1.38 [5, 7, 14, 15, 17].

Considering also the rarity of the disease, the identified variants have a poor clinical use in predicting the individual risk, especially if considering the general population. A possible approach to improve usefulness of genetic risk markers could be to combine the SNPs in a polygenic risk score (PRS) in order to have a better estimation of their cumulative effect on the risk of developing the disease. This method has been successfully applied to several diseases including breast, prostate, colorectal, and pancreatic cancer [22–28]. For myeloma, a PRS was briefly mentioned in the latest GWAS publication [17]. An earlier study compared a 16-SNP PRS in familial and sporadic MM cases [29]. A PRS including all the known risk SNPs has been also evaluated in African-Americans [30].

<sup>1</sup>Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany. <sup>2</sup>Department of Biology, University of Pisa, Pisa, Italy. <sup>3</sup>Department of Hematology, University Hospital of Cracow, Cracow, Poland. <sup>4</sup>Department of Hematology, Aarhus University Hospital, Aarhus, Denmark. <sup>5</sup>Hematology Division, Chaim Sheba Medical Center, Tel Hashomer, Israel. <sup>6</sup>Department of Hematology, Sea Hospital, Gdynia, Poland. <sup>7</sup>Department of Internal and Occupational Diseases, Hypertension and Clinical Oncology, Medical University Wrocław, Wrocław, Poland. <sup>8</sup>Cancer Research Center of Lyon/Hospices Civils de Lyon, Lyon, France. <sup>9</sup>Hematology Clinic, Holycross Cancer Center, Kielce, Poland. <sup>10</sup>Department of Hematology, Institute of Hematology and Transfusion Medicine, Warsaw, Poland. <sup>11</sup>Genomic Oncology Area, GENYO. Centre for Genomics and Oncological Research: Pfizer, University of Granada/Andalusian Regional Government, Granada, Spain. <sup>12</sup>Hematology department, Virgen de las Nieves University Hospital, Granada, Spain. <sup>13</sup>Third Department of Internal Medicine, Semmelweis University, Budapest, Hungary. <sup>14</sup>Department of Internal and Occupational Diseases, Medical University Wrocław, Wrocław, Poland. <sup>15</sup>Department of Hematology, Odense University Hospital, Odense, Denmark. <sup>16</sup>Institut Curie, PSL Research University, Mines ParisTech Inserm, U900 Paris, France. <sup>17</sup>Hematology Department, Teaching Hospital No 1, Rzeszów, Poland. <sup>18</sup>Department of Hematology, Rigshospitalet, Copenhagen University, Copenhagen, Denmark. <sup>19</sup>Clinical and Experimental Medicine, Section of Hematology, University of Pisa, Pisa, Italy. <sup>20</sup>Department of Haematology, Military Institute of Medicine, Warsaw, Poland. <sup>21</sup>Department of Hematology, Rydygier Specialistic Hospital, Cracow, Poland. <sup>22</sup>Life and Health Sciences Research Institute (ICVS), School of Health Sciences/Molecular Oncology Research Center, University of Minho, Braga, Portugal. <sup>23</sup>Department of Hematology, Jagiellonian University Medical College, Cracow, Poland. <sup>24</sup>Maria Skłodowska-Curie National Research Institute of Oncology, Warsaw, Poland. <sup>25</sup>National Research Centre for the Working Environment, DK-2100 Copenhagen, Denmark. <sup>26</sup>Institute of Molecular Medicine, University of Southern Denmark, Odense, Denmark. <sup>27</sup>ICVS/3B's - PT Government Associate Laboratory, Braga/Guimarães, Portugal. <sup>28</sup>Molecular Oncology Research Center, Barretos Cancer Hospital, S.Paulo, Brazil. <sup>29</sup>Unité de Génomique du Myélome, Institut Universitaire du Cancer Toulouse – Oncopole, Toulouse, France. <sup>30</sup>Department of Hematology, University Hospital Bydgoszcz, Bydgoszcz, Poland. <sup>31</sup>Medical University of Lublin, Lublin, Poland. <sup>32</sup>Department of Hematology, N. Copernicus Town Hospital, Torun, Poland. <sup>33</sup>CRCINA, INSERM, CNRS, Université d'Angers, Université de Nantes, Nantes, France. ✉email: f.canzian@dkfz.de

Received: 25 March 2020 Revised: 4 October 2021 Accepted: 8 October 2021

Published online: 30 November 2021

The aim of this work is to use the International Multiple Myeloma (IMMENSE) consortium to establish a PRS for MM and provide an evaluation of the PRS performance in an independent set of MM cases and controls.

## MATERIALS AND METHODS

### Study population

We used DNA samples from 2361 MM patients and 1415 controls from 7 countries (Denmark, France, Hungary, Israel, Italy, Poland, and Portugal) within the IMMENSE consortium [6], for whom information on sex and age was available. Cases were defined by a confirmed diagnosis of MM according to the International Myeloma Working Group criteria [31]. Controls were selected from the general population, from hospitalized subjects with different diagnoses excluding cancer, or from blood donors. Characteristics of the study population are summarized in Table 1.

### SNP selection

To build the PRS we used 23 SNPs shown to be associated with MM risk at genome-wide significance level ( $p < 5 \times 10^{-8}$ ) by previous GWAS [5, 7, 14, 15, 17]. We did not include variants reported to be associated with MM risk but not at genome-wide level of significance (e.g., those reported by Erickson et al. [9]). Characteristics of the SNPs included in the PRS are summarized in Supplementary Table 1.

### Genotyping and PRS computation

Genotyping was performed using TaqMan technology (ThermoFisher Applied Biosystems, Waltham MA, USA) according to the manufacturer's recommendations. TaqMan assays were not available for some SNPs, therefore we replaced them with surrogates in high linkage disequilibrium ( $r^2 > 0.9$ ), as detailed in Supplementary Table 1.

For each SNP, the number of alleles associated with higher MM risk were counted and added up for each study subject, resulting in an unweighted PRS, which had a theoretical range from 0 (no MM risk alleles) to 46 (all risk alleles are present at each SNP in homozygosity). In addition, we built a weighted PRS by using the ORs of the codominant model of the association of each variant with MM risk in the IMMENSE population as coefficients to weight the relative effects of the risk SNPs. For each SNP in the weighted PRS, a value of 0 was assigned if 0 risk alleles were present, the  $\ln(\text{OR})$  of the heterozygous was assigned if one risk allele was present, and the  $\ln(\text{OR})$  of the homozygous was assigned if two risk alleles were present. Then all the values were summed among them for each subject. We built alternative weighted PRSs by using ORs from the literature, or values calculated in our dataset. Only a subset of the study subjects (1426 cases and 969 controls) had a 100% SNP call rate. Therefore, in order to be able to compute comparable score values for all study subjects, we also considered "scaled" scores, in which the PRS values for each subject were multiplied by the ratio between 23 (total number of SNPs) and the number of effectively genotyped SNPs for the subject in question. For both

PRSs (weighted and unweighted), we calculated quintiles based on the distribution of values in the controls.

The formulas for the unweighted and weighted scores are respectively  $\sum_1^m aj$  and  $\sum_1^m aXj$ , where  $a$  = number of risk alleles (0, 1, 2),  $m$  = total number of SNPs (23),  $j$  =  $j$ th subject,  $X$  =  $\ln(\text{OR})$ . Supplementary Table 2 shows an example of how the scores were generated.

### Data filtering and statistical analysis

Samples with call rate less than 80% were not included in subsequent analysis. Pearson chi square was used to test departure from Hardy–Weinberg equilibrium (HWE) in the overall control group and in the individual countries.

To validate the associations between the individual SNPs and MM risk, we used logistic regression according to the log-additive and codominant models, using the more common allele in controls as the reference category.

We analyzed the association between the PRSs and MM risk by logistic regression. Age-stratified analyses were performed by comparing all controls with younger or older cases, with cutpoints at 55 (to distinguish between early onset and non-early onset cases), 61 (median age at onset of the cases in this study), or 69 years of age (median age at onset of MM, <https://seer.cancer.gov/statfacts/html/mulmy.html>) [32]. All analyses were adjusted for age, sex, and geographic region of origin.

We set up receiver operating characteristic (ROC) curves and calculated the areas under the curve (AUC), to determine the performance of the PRSs in discriminating MM cases from individuals without the disease.

## RESULTS

We genotyped a total of 3376 subjects (2361 cases and 1415 controls). Controls from Portugal resulted out of HWE for SNPs rs877529 and rs4325816 in one 384-well plate (using a Bonferroni-corrected threshold of  $p < 0.002$ ). Therefore, genotypes of Portuguese subjects for those two SNPs were dropped from the dataset. The remaining data were used for further statistical analyses. Duplicated samples (8% of the total) showed a concordance rate higher than 99%.

The associations between 12 of the SNPs and MM risk were replicated in IMMENSE ( $p < 0.05$ ) (Table 2). Regardless of statistical significance, all SNPs showed ORs going in the same directions as originally reported in the literature.

We observed strong associations between the PRS and MM risk (Table 3). When we computed the association between the PRSs and MM risk considering only 1426 cases and 969 controls with a call rate of 100%, we observed an  $\text{OR} = 3.18$ , 95% CI 2.34–4.33,  $p = 1.62 \times 10^{-13}$  for the highest vs. lowest quintile of the unweighted score and  $\text{OR} = 3.44$ , 95% CI 2.53–4.69,  $p = 4.86 \times 10^{-15}$  for the highest vs. lowest quintile of the weighted score. Results were very similar when we considered the whole dataset including 2361 cases and 1415 controls and "scaled" PRSs (Table 3), as well as when we built weighted scores using ORs for each SNP from the original GWASs (Table 3).

A histogram showing the difference in number of risk alleles (unweighted PRS) between cases and controls is shown in Supplementary Fig. 1.

In order to focus on the extreme parts of the risk distribution, we also calculated the difference in risk of subjects in the 95th percentile compared to subjects in the 5th percentile, and we found a substantial difference in risk ( $\text{OR} = 5.77$ , 95% CI 2.37–14.06,  $p = 1.12 \times 10^{-4}$ ). Furthermore, we compared the subjects in the 95th percentile with subjects in the middle of the score distribution (third quintile) and we obtained an  $\text{OR} = 4.22$ , 95% CI 2.11–8.44,  $p = 4.52 \times 10^{-5}$ . All the tail distribution results are shown in Table 4.

In addition, we performed case-control analyses stratifying the cases by age at diagnosis. We used three age cutpoints: 55, 61, and 69. The PRS was associated with MM risk in all strata, without differences in risk due to age of onset (data not shown).

The AUCs for each score are shown in Table 5. The best performance was observed for the unweighted PRS when

**Table 1.** Description of the study population.

	Cases	Controls	Total
Country			
Denmark	299	478	777
France	467	176	643
Hungary	104	81	185
Israel	81	68	149
Italy	251	224	475
Poland	1034	267	1301
Portugal	125	121	246
Total	2361	1415	3776
Sex			
Male	52.6%	52.4%	52.5%
Female	47.4%	47.6%	47.5%
Median age	61	50	58

**Table 2.** Association between the selected SNPs and MM risk in the IMMENSE population.

SNP	Cases						Controls						Log-additive model						Codominant model						p trend
	MM		mm		Mm		MM		mm		Mm		M vs. m		MM vs. Mm		MM vs. mm		M vs. m		MM vs. mm				
	MM	Mm	mm	MM	Mm	mm	MM	Mm	mm	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	OR	95% CI	P	
rs6746082	1542	588	93	896	422	68	896	422	68	0.89	0.78–1.02	0.098	0.90	0.76–1.07	0.260	0.76	0.52–1.12	0.165	0.76	0.52–1.12	0.165	0.76	0.52–1.12	0.165	0.481
rs4325816	1423	690	91	818	431	79	818	431	79	0.89	0.78–1.02	0.092	0.95	0.80–1.12	0.529	<b>0.69</b>	<b>0.48–0.99</b>	<b>0.048</b>	<b>0.69</b>	<b>0.48–0.99</b>	<b>0.048</b>	<b>0.69</b>	<b>0.48–0.99</b>	<b>0.048</b>	0.197
rs1052501	1351	817	130	924	406	46	924	406	46	<b>1.28</b>	<b>1.17–1.55</b>	<b>3.1 × 10<sup>-5</sup></b>	<b>1.43</b>	<b>1.21–1.70</b>	<b>3.2 × 10<sup>-5</sup></b>	1.50	1.00–2.24	0.050	1.50	1.00–2.24	0.050	1.50	1.00–2.24	0.050	<b>1.1 × 10<sup>-4</sup></b>
rs10936599	1435	769	115	796	501	88	796	501	88	<b>0.84</b>	<b>0.73–0.95</b>	<b>0.007</b>	<b>0.83</b>	<b>0.71–0.99</b>	<b>0.034</b>	<b>0.70</b>	<b>0.49–0.99</b>	<b>0.044</b>	<b>0.70</b>	<b>0.49–0.99</b>	<b>0.044</b>	<b>0.70</b>	<b>0.49–0.99</b>	<b>0.044</b>	0.121
rs2548594	1326	808	157	773	511	99	773	511	99	0.94	0.82–1.06	0.298	0.97	0.82–1.14	0.679	0.83	0.60–1.13	0.237	0.83	0.60–1.13	0.237	0.83	0.60–1.13	0.237	0.213
rs6595443	682	1119	465	497	619	266	497	619	266	<b>1.15</b>	<b>1.03–1.29</b>	<b>0.011</b>	<b>1.29</b>	<b>1.08–1.55</b>	<b>0.004</b>	<b>1.29</b>	<b>1.03–1.60</b>	<b>0.027</b>	<b>1.29</b>	<b>1.03–1.60</b>	<b>0.027</b>	<b>1.29</b>	<b>1.03–1.60</b>	<b>0.027</b>	<b>0.017</b>
rs34229995	2136	134	10	1339	63	2	1339	63	2	1.32	0.95–1.83	0.102	1.25	0.87–1.80	0.230	2.71	0.54–13.54	0.225	2.71	0.54–13.54	0.225	2.71	0.54–13.54	0.225	0.434
rs2285803	975	981	292	666	598	134	666	598	134	<b>1.15</b>	<b>1.02–1.29</b>	<b>0.022</b>	1.08	0.91–1.28	0.365	<b>1.40</b>	<b>1.07–1.83</b>	<b>0.013</b>	<b>1.40</b>	<b>1.07–1.83</b>	<b>0.013</b>	<b>1.40</b>	<b>1.07–1.83</b>	<b>0.013</b>	<b>0.007</b>
rs9373839	1492	734	93	881	388	49	881	388	49	1.03	0.90–1.19	0.639	1.09	0.92–1.30	0.322	0.91	0.60–1.37	0.640	0.91	0.60–1.37	0.640	0.91	0.60–1.37	0.640	0.767
rs4487645	1280	871	152	614	625	158	614	625	158	<b>0.72</b>	<b>0.64–0.82</b>	<b>1.7 × 10<sup>-7</sup></b>	<b>0.74</b>	<b>0.63–0.87</b>	<b>3.6 × 10<sup>-5</sup></b>	<b>0.51</b>	<b>0.38–0.68</b>	<b>3.3 × 10<sup>-6</sup></b>	<b>0.51</b>	<b>0.38–0.68</b>	<b>3.3 × 10<sup>-6</sup></b>	<b>0.51</b>	<b>0.38–0.68</b>	<b>3.3 × 10<sup>-6</sup></b>	<b>2.4 × 10<sup>-6</sup></b>
rs17507636	1315	813	144	715	540	136	715	540	136	<b>0.78</b>	<b>0.69–0.88</b>	<b>6.2 × 10<sup>-5</sup></b>	<b>0.79</b>	<b>0.67–0.93</b>	<b>0.005</b>	<b>0.59</b>	<b>0.44–0.80</b>	<b>0.001</b>	<b>0.59</b>	<b>0.44–0.80</b>	<b>0.001</b>	<b>0.59</b>	<b>0.44–0.80</b>	<b>0.001</b>	<b>0.001</b>
rs2170352	1254	847	182	770	538	85	770	538	85	1.08	0.96–1.23	0.210	1.02	0.86–1.20	0.839	1.31	0.96–1.82	0.091	1.31	0.96–1.82	0.091	1.31	0.96–1.82	0.091	0.431
rs7781265	1816	474	40	1075	269	18	1075	269	18	1.08	0.91–1.28	0.383	1.05	0.87–1.28	0.601	1.35	0.70–2.59	0.373	1.35	0.70–2.59	0.373	1.35	0.70–2.59	0.373	0.637
rs1948915	973	1048	287	639	611	150	639	611	150	<b>1.60</b>	<b>1.03–1.30</b>	<b>0.012</b>	1.13	0.96–1.33	0.151	<b>1.38</b>	<b>0.96–1.33</b>	<b>0.014</b>	<b>1.38</b>	<b>0.96–1.33</b>	<b>0.014</b>	<b>1.38</b>	<b>0.96–1.33</b>	<b>0.014</b>	<b>0.011</b>
rs2811710	1085	951	261	548	578	199	548	578	199	<b>0.83</b>	<b>0.74–0.94</b>	<b>0.002</b>	0.86	0.73–1.02	0.092	<b>0.68</b>	<b>0.53–0.87</b>	<b>0.002</b>	<b>0.68</b>	<b>0.53–0.87</b>	<b>0.002</b>	<b>0.68</b>	<b>0.53–0.87</b>	<b>0.002</b>	0.117
rs7187359	1108	905	248	733	534	129	733	534	129	1.08	0.96–1.21	0.204	1.05	0.89–1.24	0.593	1.21	0.92–1.59	0.175	1.21	0.92–1.59	0.175	1.21	0.92–1.59	0.175	0.137
rs2790454	1310	836	157	768	527	103	768	527	103	0.92	0.81–1.04	0.178	0.95	0.80–1.12	0.522	0.80	0.58–1.09	0.152	0.80	0.58–1.09	0.152	0.80	0.58–1.09	0.152	0.201
rs7193541	884	1069	367	480	644	256	480	644	256	0.93	0.83–1.03	0.177	0.87	0.73–1.04	0.124	0.88	0.70–1.10	0.271	0.88	0.70–1.10	0.271	0.88	0.70–1.10	0.271	0.722
rs4273077	1626	554	53	1042	260	24	1042	260	24	<b>1.26</b>	<b>1.07–1.49</b>	<b>0.006</b>	<b>1.29</b>	<b>1.06–1.57</b>	<b>0.010</b>	1.41	0.81–2.48	0.227	1.41	0.81–2.48	0.227	1.41	0.81–2.48	0.227	0.371
rs11086029	1336	754	107	854	450	62	854	450	62	1.14	1.00–1.31	0.052	1.15	0.97–1.36	0.116	1.31	0.89–1.91	0.167	1.31	0.89–1.91	0.167	1.31	0.89–1.91	0.167	0.217
rs6066835	1924	374	44	1159	229	7	1159	229	7	1.15	0.95–1.39	0.162	1.07	0.87–1.32	0.500	<b>2.86</b>	<b>1.05–7.80</b>	<b>0.040</b>	<b>2.86</b>	<b>1.05–7.80</b>	<b>0.040</b>	<b>2.86</b>	<b>1.05–7.80</b>	<b>0.040</b>	0.884
rs138745	970	1058	296	597	619	172	597	619	172	1.10	0.98–1.23	0.107	1.08	0.91–1.28	0.372	1.23	0.95–1.57	0.111	1.23	0.95–1.57	0.111	1.23	0.95–1.57	0.111	0.397
rs877529	678	986	486	471	565	254	471	565	254	<b>1.22</b>	<b>1.09–1.37</b>	<b>3.8 × 10<sup>-4</sup></b>	1.21	1.01–1.46	0.440	<b>1.50</b>	<b>1.20–1.88</b>	<b>4.1 × 10<sup>-4</sup></b>	<b>1.50</b>	<b>1.20–1.88</b>	<b>4.1 × 10<sup>-4</sup></b>	<b>1.50</b>	<b>1.20–1.88</b>	<b>4.1 × 10<sup>-4</sup></b>	<b>1.4 × 10<sup>-4</sup></b>

All analyses were adjusted for age, sex, and geographic region of origin. Results in bold are statistically significant ( $p < 0.05$ ).

M major allele, m minor allele, OR odds ratio, CI confidence interval, as calculated in IMMENSE.

**Table 3.** Associations between PRSs and MM risk with the different types of scores.

Type of score	Quintiles	OR <sup>a</sup>	95% CI <sup>a</sup>	P <sub>value</sub>
Unweighted, subjects with 100% call rate	1	1.00	–	Ref.
	2	0.63	0.46–0.86	0.004
	3	3.16	2.31–4.31	$4.33 \times 10^{-13}$
	4	2.42	1.81–3.24	$3.17 \times 10^{-9}$
	5	3.18	2.34–4.33	$1.62 \times 10^{-13}$
	Continuous <sup>b</sup>	1.43	1.34–1.54	$7.00 \times 10^{-23}$
Unweighted scaled, all subjects	1	1.00	–	Ref.
	2	1.52	1.17–1.97	0.002
	3	1.44	1.13–1.83	0.003
	4	2.20	1.73–2.80	$1.45 \times 10^{-10}$
	5	2.93	2.28–3.78	$9.00 \times 10^{-16}$
	Continuous <sup>b</sup>	1.29	1.22–1.37	$1.00 \times 10^{-17}$
Weighted, subjects with 100% call rate <sup>c</sup>	1	1.00	–	Ref.
	2	1.33	0.95–1.86	0.096
	3	1.60	1.15–2.23	0.005
	4	2.43	1.77–3.35	$4.78 \times 10^{-8}$
	5	3.44	2.53–4.69	$3.55 \times 10^{-15}$
	Continuous <sup>b</sup>	1.37	1.28–1.46	$2.00 \times 10^{-18}$
Weighted scaled, all subjects <sup>c</sup>	1	1.00	–	Ref.
	2	1.29	0.98–1.70	0.068
	3	1.53	1.17–2.01	0.002
	4	2.24	1.72–2.91	$1.68 \times 10^{-9}$
	5	3.12	2.42–4.02	$2.00 \times 10^{-17}$
	Continuous <sup>b</sup>	1.33	1.26–1.41	$3.00 \times 10^{-22}$
Weighted 100% call rate using GWAS OR <sup>d</sup>	1	1.00	–	Ref.
	2	1.18	0.84–1.65	0.334
	3	1.56	1.12–2.17	0.008
	4	2.17	1.59–2.97	$1.29 \times 10^{-6}$
	5	3.24	2.39–4.39	$3.93 \times 10^{-14}$
	Continuous <sup>b</sup>	1.35	1.27–1.45	$2.00 \times 10^{-17}$
Weighted scaled using GWAS OR <sup>d</sup>	1	1.00	–	Ref.
	2	1.21	0.93–1.60	0.161
	3	1.56	1.20–2.04	0.001
	4	2.02	1.57–2.62	$7.86 \times 10^{-8}$
	5	2.89	2.25–3.71	$9.00 \times 10^{-16}$
	Continuous <sup>b</sup>	1.31	1.24–1.38	$9.00 \times 10^{-20}$

<sup>a</sup>OR odds ratio; CI confidence interval; all analyses were adjusted for age, sex and geographic region of origin.

<sup>b</sup>The unit for the analysis with the continuous variable was the increment of one quintile.

<sup>c</sup>The weights used to build this score were the ORs of the associations between the individual SNPs and MM risk observed in the IMMEnSE population.

<sup>d</sup>The weights used to build this score were the ORs of the associations between the individual SNPs and MM risk observed in the literature.

considering only subjects with 100% call rate (AUC = 0.64, 95% CI = 0.62–0.67).

## DISCUSSION

Twenty-three SNPs affecting risk of MM were identified through GWAS. Since individually they do not explain a large proportion of the disease risk, we combined them in a PRS, which showed association with MM risk with strong statistical significance. Our results are encouraging, since when comparing the tails of the PRS distribution we observed a fourfold or more increase in risk.

The best area under the curve associated with the PRS was modest (AUC = 0.64, 95% CI = 0.62–0.67). However, this test could show a much better predictive ability in a selected population at

already increased risk, such as individuals with MGUS or SM patients. We expect that the PRS performance will improve as more variants associated with MM are discovered, as shown by studies on other cancer types [23, 26, 27]. A further step to the clinical use of PRS is to combine them with environmental or lifestyle risk factors, as well as family history. We can envisage that in the middle/long term an enhanced MM risk PRS could become a powerful prediction tool for individualized risk stratification. Genotyping of risk loci will be done quickly and inexpensively in large groups of the population. Information on risk loci will be combined with questionnaire data on non-genetic risk factors, and specialized algorithms will estimate disease risk in a personalized manner. This will allow to adopt preventive measures, such as enhanced surveillance or intensified screening of people at high risk.

**Table 4.** Associations between subjects in the 95<sup>th</sup> percentile vs 5<sup>th</sup> and third quintile and MM risk with the different types of scores.

Type of score	No of cases	No of controls	Distribution	OR <sup>a</sup>	95% CI <sup>a</sup>	P <sub>value</sub>
Unweighted 100% call rate	202	44	95% vs 5%	5.77	2.37–14.06	$1.12 \times 10^{-4}$
	476	142	95% vs third quintile	4.22	2.11–8.44	$4.52 \times 10^{-5}$
Unweighted scaled	356	141	95% vs 5%	4.12	2.42–7.01	$1.81 \times 10^{-7}$
	745	407	95% vs third quintile	3.05	2.15–4.32	$3.73 \times 10^{-10}$
Weighted 100% call rate	221	97	95% vs 5%	6.81	3.52–13.16	$1.20 \times 10^{-8}$
	398	241	95% vs third quintile	3.05	1.98–4.70	$4.41 \times 10^{-7}$
Weighted scaled	316	141	95% vs 5%	4.29	2.52–7.30	$7.95 \times 10^{-8}$
	646	352	95% vs third quintile	2.41	1.68–3.45	$1.64 \times 10^{-6}$

**Table 5.** Areas under the curve (AUC) for each PRS.

	AUC	95% CI
Unweighted score		
Subjects with call rate = 100%	0.644	0.622–0.666
“Scaled” score, all subjects	0.601	0.583–0.619
Weighted score calculated using ORs estimated in IMMENSE		
Subjects with call rate = 100%	0.628	0.605–0.650
“Scaled” score, all subjects	0.615	0.597–0.633
Weighted score calculated using ORs from published GWAS		
Subjects with call rate = 100%	0.628	0.606–0.650
“Scaled” score, all subjects	0.609	0.591–0.627

A limitation of this work is that the individuals used are all of European origin, making it difficult to generalize the data for other ethnicities. The same PRS was recently studied in African-Americans, with results comparable to those of European descent people [30]. Another limitation is that we examined only genetic polymorphisms. It would be worth exploring whether a multifactorial score including also non-genetic risk factors could have a better predictive power. Unfortunately, we do not have complete data about known MM risk factors in IMMENSE, therefore we can not explore multifactorial risk scores with meaningful numbers of cases and controls.

In conclusion, we found a convincing association of a 23-SNP PRS and MM risk. Our work provides additional validation of previously discovered MM risk variants and of their combination into a PRS, which is a first step toward the use of genetic background in the prevention of the disease. Additional risk SNP discovery will allow to generate PRS with a better accuracy and a clearer usefulness.

## DATA AVAILABILITY

The dataset underlying this manuscript has been submitted to the European Genome-phenome Archive (EGA) under accession number EGAS00001005654.

## REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2018;68:394–424.
- Kyle RA, Therneau TM, Rajkumar SV, Larson DR, Plevak MF, Offord JR, et al. Prevalence of monoclonal gammopathy of undetermined significance. *N Engl J Med.* 2006;354:1362–9.
- Kyle RA, Larson DR, Therneau TM, Dispenzieri A, Kumar S, Cerhan JR, et al. Long-term follow-up of monoclonal gammopathy of undetermined significance. *N Engl J Med.* 2018;378:241–9.
- Agarwal A, Ghobrial IM. Monoclonal gammopathy of undetermined significance and smoldering multiple myeloma: a review of the current understanding of epidemiology, biology, risk stratification, and management of myeloma precursor disease. *Clin Cancer Res.* 2013;19:985–94.

- Broderick P, Chubb D, Johnson DC, Weinhold N, Försti A, Lloyd A, et al. Common variation at 3p22.1 and 7p15.3 influences multiple myeloma risk. *Nat Genet.* 2011;44:58–61.
- Martino A, Sainz J, Buda G, Jamrozik K, Reis RM, García-Sanz R, et al. Genetics and molecular epidemiology of multiple myeloma: the rationale for the IMMENSE consortium (review). *Int J Oncol.* 2012;40:625–38.
- Chubb D, Weinhold N, Broderick P, Chen B, Johnson DC, Försti A, et al. Common variation at 3q26.2, 6p21.33, 17p11.2 and 22q13.1 influences multiple myeloma risk. *Nat Genet.* 2013;45:1221–5.
- Weinhold N, Johnson DC, Chubb D, Chen B, Försti A, Hosking FJ, et al. The CCND1 c.870G>A polymorphism is a risk factor for t(11;14)(q13;q32) multiple myeloma. *Nat Genet.* 2013;45:522–5.
- Erickson SW, Raj VR, Stephens OW, Dhakal I, Chavan SS, Sanathkumar N, et al. Genome-wide scan identifies variant in 2q12.3 associated with risk for multiple myeloma. *Blood.* 2014;124:2001–3.
- Campa D, Martino A, Varkonyi J, Lesueur F, Jamrozik K, Landi S, et al. Risk of multiple myeloma is associated with polymorphisms within telomerase genes and telomere length. *Int J Cancer.* 2015;136:E351–358.
- Campa D, Martino A, Macaudo A, Dudziński M, Suska A, Druzd-Sitek A, et al. Genetic polymorphisms in genes of class switch recombination and multiple myeloma risk and survival: an IMMENSE study. *Leuk Lymphoma.* 2019;60:1803–11.
- Macaudo A, Castelli E, Buda G, Pelosini M, Butrym A, Watek M, et al. Inherited variation in the xenobiotic transporter pathway and survival of multiple myeloma patients. *Br J Haematol.* 2018;183:375–84.
- Martino A, Campa D, Buda G, Sainz J, García-Sanz R, Jamrozik K, et al. Polymorphisms in xenobiotic transporters ABCB1, ABCG2, ABCC2, ABCC1, ABCC3 and multiple myeloma risk: a case-control study in the context of the International Multiple Myeloma rESEarch (IMMENSE) consortium. *Leukemia.* 2012;26:1419–22.
- Swaminathan B, Thorleifsson G, Jöud M, Ali M, Johnsson E, Ajore R, et al. Variants in ELL2 influencing immunoglobulin levels associate with multiple myeloma. *Nat Commun.* 2015;6:7213.
- Mitchell JS, Li N, Weinhold N, Försti A, Ali M, van Duin M, et al. Genome-wide association study identifies multiple susceptibility loci for multiple myeloma. *Nat Commun.* 2016;7:12050.
- Halvarsson B-M, Wihlborg A-K, Ali M, Lemonakis K, Johnsson E, Niroula A, et al. Direct evidence for a polygenic etiology in familial multiple myeloma. *Blood Adv.* 2017;1:619–23.
- Went M, Sud A, Försti A, Halvarsson B-M, Weinhold N, Kimber S, et al. Identification of multiple risk loci and regulatory mechanisms influencing susceptibility to multiple myeloma. *Nat Commun.* 2018;9:3707.
- Pertesi M, Vallée M, Wei X, Revuelta MV, Galia P, Demangel D, et al. Exome sequencing identifies germline variants in DIS3 in familial multiple myeloma. *Leukemia.* 2019;33:2324–30.
- Wei X, Calvo-Vidal MN, Chen S, Wu G, Revuelta MV, Sun J, et al. Germline lysine-specific demethylase 1 (LSD1/KDM1A) mutations confer susceptibility to multiple myeloma. *Cancer Res.* 2018;78:2747–59.
- Went M, Sud A, Speedy H, Sunter NJ, Försti A, Law PJ, et al. Genetic correlation between multiple myeloma and chronic lymphocytic leukaemia provides evidence for shared aetiology. *Blood Cancer J.* 2018;9:1.
- Pertesi M, Went M, Hansson M, Hemminki K, Houlston RS, Nilsson B. Genetic predisposition for multiple myeloma. *Leukemia.* 2020;34:697–708.
- Burton H, Chowdhury S, Dent T, Hall A, Pashayan N, Pharoah P. Public health implications from COGS and potential for risk stratification and screening. *Nat Genet.* 2013;45:349–51.
- Hsu L, Jeon J, Brenner H, Gruber SB, Schoen RE, Berndt SI, et al. A model to determine colorectal cancer risk using common genetic susceptibility loci. *Gastroenterology.* 2015;148:1330–9.e14.

24. Lecarpentier J, Silvestri V, Kuchenbaecker KB, Barrowdale D, Dennis J, McGuffog L, et al. Prediction of breast and prostate cancer risks in male BRCA1 and BRCA2 mutation carriers using polygenic risk scores. *J Clin Oncol.* 2017;35:2240–50.
25. Torkamani A, Wineinger NE, Topol EJ. The personal and clinical utility of polygenic risk scores. *Nat Rev Genet.* 2018;19:581–90.
26. Seibert TM, Fan CC, Wang Y, Zuber V, Karunamuni R, Parsons JK, et al. Polygenic hazard score to guide screening for aggressive prostate cancer: development and validation in large scale cohorts. *BMJ.* 2018;360:j5757.
27. Mavaddat N, Michailidou K, Dennis J, Lush M, Fachal L, Lee A, et al. Polygenic risk scores for prediction of breast cancer and breast cancer subtypes. *Am J Hum Genet.* 2019;104:21–34.
28. Galeotti AA, Gentiluomo M, Rizzato C, Obazee O, Neoptolemos JP, Pasquali C, et al. Polygenic and multifactorial scores for pancreatic ductal adenocarcinoma risk prediction. *J Med Genet.* 2021;58:369–77.
29. Halvarsson B-M, Wihlborg A-K, Ali M, Lemonakis K, Johnsson E, Niroula A, et al. Direct evidence for a polygenic etiology in familial multiple myeloma. *Blood Adv.* 2017;1:619–23.
30. Du Z, Weinhold N, Song GC, Rand KA, Van Den Berg DJ, Hwang AE, et al. A meta-analysis of genome-wide association studies of multiple myeloma among men and women of African ancestry. *Blood Adv.* 2020;4:181–90.
31. Rajkumar SV. Updated diagnostic criteria and staging system for multiple myeloma. *Am Soc Clin Oncol Educ Book.* 2016;35:e418–23.
32. Kumar SK, Rajkumar V, Kyle RA, van Duin M, Sonneveld P, Mateos M-V, et al. Multiple myeloma. *Nat Rev Dis Prim.* 2017;3:17046.

## ACKNOWLEDGEMENTS

This work was partially supported by intramural funds of the University of Pisa, DKFZ, and University Hospital of Southern Jutland, Denmark, and by a grant of the French National Cancer Institute (INCA). The authors wish to thank Dr. Dominic Edelmann (Division of Biostatistics, DKFZ) for helpful advice about data analysis.

## AUTHOR CONTRIBUTIONS

DC and FC conceived and designed the study. CP performed labwork. CP, FC, AM, and DC performed data quality control and statistical analyses. FC, AM, and DC drafted the manuscript. All other authors provided samples and data. All authors critically read, commented, and approved the manuscript.

## FUNDING

Open Access funding enabled and organized by Projekt DEAL.

## COMPETING INTERESTS

The authors declare no competing interests.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

The IMMEnSE study protocol was approved by the Ethics Committee of the Medical Faculty of the University of Heidelberg (reference number: S-004/2020). Following the guidelines of the Declaration of Helsinki, written informed consent was obtained from each participant.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41431-021-00986-8>.

**Correspondence** and requests for materials should be addressed to Federico Canzian.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021