

1 **Prediction of Homologous Recombination Deficiency from OncoPrint Comprehensive**

2 **Assay Plus Correlating with SOPHiA DDM HRD Solution**

3

4 Jun Kang<sup>1</sup>, Kiyong Na<sup>2</sup>, Haeyoun Kang<sup>3</sup>, Uiju Cho<sup>4</sup>, Sun Young Kwon<sup>5</sup>, Sohyun Hwang<sup>6</sup>,  
5 Ahwon Lee<sup>1,7</sup>, The Molecular Pathology Study Group of Korean Society of Pathologists

6

7 <sup>1</sup>Department of Hospital Pathology, Seoul St. Mary's Hospital, College of Medicine, The  
8 Catholic University of Korea, Seoul, Korea

9 <sup>2</sup>Department of Pathology, Kyung Hee University College of Medicine, Kyung Hee  
10 University Hospital, Seoul, Republic of Korea

11 <sup>3</sup>Department of Pathology, CHA Bundang Medical Center, CHA University, Seongnam,  
12 Korea

13 <sup>4</sup>Department of Pathology, St. Vincent's Hospital, College of Medicine, The Catholic  
14 University of Korea, Seoul, Korea

15 <sup>5</sup>Department of pathology, Dongsan Hospital, School of Medicine, Keimyung University

16 <sup>6</sup>CHA Future Medicine Research Institute, CHA Bundang Medical Center, Seongnam, Korea

17 <sup>7</sup>Cancer Research Institute, The Catholic University of Korea, Seoul, Korea

18

19 Short title: Prediction of HRD

20

21 Corresponding author

22 Ahwon Lee

23 Professor

24 Department of Hospital Pathology

25 Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea

26 222, Banpo-daero, Seocho-gu, Seoul, 06591, Republic of Korea

27 Phone: +82-2-2258-1626

28 Fax: +82-2-2258-1627

29 E-mail: [klee@catholic.ac.kr](mailto:klee@catholic.ac.kr)

30

31 **ABSTRACT**

32 **Objective** Poly(ADP-ribose) polymerase (PARP) inhibitors are used for targeted therapy for  
33 ovarian cancer with homologous recombination deficiency (HRD). In this study, we aimed to  
34 develop a homologous recombination deficiency prediction model to predict the genomic  
35 integrity (GI) index of the SOPHiA DDM HRD Solution from the Oncomine Comprehensive  
36 Assay (OCA) Plus. We also tried to find a cut-off value of the genomic instability metric  
37 (GIM) of the OCA Plus that correlates with the GI index of the SOPHiA DDM HRD Solution.

38 **Methods** We included 87 cases with high-grade ovarian serous carcinoma from five tertiary  
39 referral hospitals in Republic of Korea. We developed an HRD prediction model to predict  
40 the GI index of the SOPHiA DDM HRD Solution. As predictor variables in the model, we  
41 used the HRD score, which included percent loss of heterozygosity (%LOH), percent  
42 telomeric allelic imbalance (%TAI), percent large-scale state transitions (%LST), and the  
43 genomic instability metric (GIM), provided by the OCA Plus. To build the model, we  
44 employed a penalized logistic regression technique.

45 **Results** The final model equation is  $-21.77 + 0.200 \times \text{GIM} + 0.102 \times \% \text{LOH} + 0.037 \times \% \text{TAI}$   
46  $+ 0.261 \times \% \text{LST}$ . To improve the performance of the prediction model, we added a borderline  
47 result category to the GI results. Cases with predicted values between -3 and 3 were classified  
48 as borderline. The accuracy of our HRD status prediction model was 0.947 for the training set  
49 and 0.958 for the test set. The accuracy of HRD status using GIM with a cut-off value of 16  
50 was 0.911.

51 **Conclusions** The Oncomine Comprehensive Assay Plus provides a reliable biomarker for  
52 homologous recombination deficiency.

53

54 **What is already known on this topic**

55 The OncoPrint Comprehensive Assay Plus is a targeted next-generation sequencing assay  
56 designed to detect genetic alterations in solid tumors. It has not been validated as a biomarker  
57 for PARP inhibitor response through clinical trials or a concordance test with a Food and  
58 Drug Administration–approved homologous recombination deficiency test.

59 **What this study adds**

60 This study introduces a predictive model for homologous recombination deficiency using  
61 data from the OncoPrint Comprehensive Assay Plus, which correlates with the SOPHiA  
62 DDM HRD Solution. The study provides evidence that the OncoPrint Comprehensive Assay  
63 Plus is a reliable biomarker for homologous recombination deficiency.

64 **How this study might affect research, practice or policy**

65 The strong agreement between the OncoPrint Comprehensive Assay Plus and the SOPHiA  
66 DDM HRD Solution supports its potential as a biomarker for predicting PARP inhibitor  
67 response. This finding could have implications for PARP inhibitor-related research, clinical  
68 practice, and regarding their use in clinical trials.

69

70

71

72

73

## 74 INTRODUCTION

75 Homologous recombination repair (HRR) is a DNA repair mechanism that restores DNA  
76 double-strand breaks (DSBs) in cells. This mechanism is essential for maintaining genomic  
77 stability and preventing the accumulation of DNA damage that can lead to mutations and  
78 other genetic alterations. When HRR is impaired, such as through mutations in genes  
79 involved in this repair pathway, it can lead to a condition known as homologous  
80 recombination deficiency (HRD). HRD has been found to be associated with an increased  
81 risk of developing certain types of cancer, including ovarian, breast, and prostate cancer<sup>1,2</sup>.  
82 The genes most commonly associated with HRD are *BRCA1* and *BRCA2*, which are tumor  
83 suppressor genes that play a critical role in HRR<sup>3</sup>. Mutations in these genes can impair the  
84 HRR pathway, leading to an increased risk of developing cancer. Other genes involved in  
85 HRR, such as *PALB2* and *RAD51*, have also been linked to HRD and an increased cancer  
86 risk<sup>4,5</sup>.

87 Poly(ADP-ribose) polymerase (PARP) inhibitors such as olaparib and talazoparib are  
88 a type of targeted therapy that work by inhibiting the function of poly(ADP-ribose)  
89 polymerase 1 (PARP1), an enzyme that is involved in the repair of single-strand DNA breaks  
90 (SSBs). PARP inhibitors have shown clinical efficacy in *BRCA1/2* mutant ovarian cancer,  
91 breast cancer, and prostate cancer<sup>6-9</sup>. These drugs have demonstrated promising results in  
92 clinical trials and have been approved by regulatory agencies for the treatment of certain  
93 types of cancer.

94 HRD can lead to abnormal DSB repair and result in genomic scars, which are large-  
95 scale genomic alterations that can be quantified by counting the number of occurrences.  
96 There are several types of genomic scars associated with HRD, including large-scale loss of  
97 heterozygosity (LOH)<sup>10</sup>, telomere allelic imbalance (TAI)<sup>11</sup>, and large-scale state transitions

98 (LST)<sup>12</sup>. The HRD score is a quantification of these genomic scars and is used to identify  
99 patients who may benefit from treatment with PARP inhibitors<sup>13 14</sup>. The HRD score is  
100 calculated based on the occurrence of these genomic scars. These tests, such as the Myriad  
101 myChoice CDx and FoundationOne CDx tests, have been approved by regulatory agencies as  
102 companion diagnostics for PARP inhibitor treatment in patients with ovarian and prostate  
103 cancer<sup>6 8 15 16</sup>.

104 The SOPHiA DDM HRD Solution is a HRD test that identifies HRR mutations  
105 through targeted sequencing and measures genomic instability (GI) through a combination of  
106 low-pass whole-genome sequencing and a deep-learning algorithm<sup>17</sup>. The GI index is a  
107 measure of genomic stability of the SOPHiA DDM HRD Solution. This index is based on the  
108 analysis of the genome-wide patterns of copy number variations (CNVs) and is used to  
109 determine the level of GI in a tumor sample. A high GI index indicates a high level of GI and  
110 is associated with HRD tumors<sup>17</sup>.

111 The OncoPrint Comprehensive Assay (OCA) Plus is a targeted next-generation  
112 sequencing (NGS) assay designed to detect genetic alterations in solid tumors. The HRD  
113 score provided by the OCA includes (1) percent LOH (%LOH), which estimates the fraction  
114 of the genome with LOH identified using genomic segmentation; (2) percent TAI (%TAI),  
115 which estimates the fraction of the genome with allelic imbalance or unequal contribution  
116 from the two alleles in the telomeres identified using genomic segmentation; and (3) percent  
117 LST (%LST), which estimates the fraction of the genome with unequal copy numbers in  
118 adjacent segments identified using genomic segmentation. These values range from 0 to 100.  
119 The genomic instability metric (GIM) is a proprietary measurement that quantifies genomic  
120 scarring associated with HRD.

121            In this study, we aimed to develop an HRD prediction model to predict the GI index  
122 of the SOPHiA DDM HRD Solution from the OCA Plus. We also tried to find a cut-off value  
123 of the GIM of the OCA Plus that correlates with the GI index of the SOPHiA DDM HRD  
124 Solution.

125

## 126 **METHODS**

### 127 **Sample Collection**

128 We included 87 cases of high-grade ovarian serous carcinoma from five tertiary referral  
129 hospitals in Republic of Korea. All cases had been tested with OCA Plus NGS panel for  
130 clinical purpose at the hospitals where patients were treated. We excluded the cases that failed  
131 to analyze HRD scores provided by the OCA Plus. In all cases, we confirmed the clinical  
132 information and tissue diagnosis and we selected paraffin blocks for the SOPHiA DDM HRD  
133 Solution. We cut all formalin-fixed paraffin-embedded (FFPE) tissue samples to a thickness  
134 of 5  $\mu\text{m}$ . We sent 10 sections to the institution in the Republic of Korea that performs the  
135 SOPHiA DDM HRD Solution.

136

### 137 **Oncomine Comprehensive Assay Plus**

138 We extracted genomic DNA by using the Maxwell RSC DNA FFPE Kit (Promega, Madison,  
139 WI, USA) in accordance with the instructions provided by the manufacturer. We determined  
140 the DNA concentration by using the Qubit ds DNA High-Sensitive Assay kit (Thermo Fisher  
141 Scientific, Waltham, MA, USA) on the Qubit fluorometer (Thermo Fisher Scientific).

142            We performed all manual library preparation by using the OCA Plus system (Thermo

143 Fisher Scientific), following the manufacturer's instructions. We conducted the multiplex  
144 polymerase chain reaction (PCR) amplification with an approximate DNA concentration of  
145 20 ng. Prior to PCR amplification, we carried out the deamination reaction in the OCA Plus  
146 by using Uracil-DNA Glycosylase, heat labile (Thermo Fisher Scientific).

147 For sequencing, we loaded the prepared libraries onto Ion 550 Chips (Thermo Fisher  
148 Scientific) according to the manufacturer's instructions and processed them using the Ion  
149 Chef System. We used the Ion S5 XL Sequencer (Thermo Fisher Scientific) for sequencing.  
150 We aligned the data to the human genome assembly 19, which served as the standard  
151 reference genome in the Ion Reporter Software (v. 5.18) (Thermo Fisher Scientific). Hospital  
152 B utilized the customized variability control informatics baseline (VCIB) for copy number  
153 analysis. The GIM was obtained from the Ion Reporter Software (v. 5.20).

154

### 155 **Genomic Instability Score Prediction Modeling**

156 We developed an HRD prediction model that aimed to predict the GI index of the SOPHiA  
157 DDM HRD Solution. The training set consisted of cases from hospital A, while the test set  
158 comprised cases from the other hospitals. The predictor variables used in the model were the  
159 HRD score, which included %LOH, %TAI, %LST, and the GIM, provided by the OCA Plus.  
160 To build the model, we employed a penalized logistic regression technique. We selected the  
161 model through repeated fivefold cross-validation on a grid of hyperparameters:  $\lambda$  ( $10^{-5}$ ,  $10^{-4}$ ,  
162  $10^{-3}$ ,  $10^{-2}$ , and  $10^{-1}$ ) and  $\alpha$  (0.0, 0.25, 0.5, 0.75, and 1.0).

163

### 164 **Assessing Model Performance**



165 We estimated the performance of the prediction based on the area under the curve of the  
166 receiver operating characteristic curve (AUROC) for the GI status and the HRD status. We  
167 considered the GI status to be positive when the GI index exceeded 0. On the other hand, we  
168 considered the HRD status to be positive if there was a BRCA1/2 pathogenic variant or if the  
169 GI status was positive. We conducted the modeling and assessment of model performance  
170 using the tidymodels and glmnet R packages. A flowchart of the study is presented in  
171 Supplemental Figure 1.

172 In accordance with the journal's guidelines, we will provide our data for independent  
173 analysis by a selected team by the Editorial Team for the purposes of additional data analysis  
174 or for the reproducibility of this study in other centers if such is requested.

175

## 176 **Research Ethics and Patient Consent**

177 The study was approved by the Institutional Review Board of the CHA Bundang Medical  
178 Center, CHA University (2023-01-010-001) and the Catholic University of Seoul Saint  
179 Mary's Hospital (KC18TNSI0361), where this study was organized.

180

## 181 **RESULTS**

### 182 **Patients**

183 The average age of the patients was 59.3 years. The majority of patients had advanced disease  
184 based on the The International Federation of Gynecology and Obstetrics (FIGO) stage.  
185 Hospital A contributed the most cases (55, accounting for 63.2% of the total). There were no  
186 significant statistical differences in patient age and the FIGO stage between the train set and

187 the test set, as shown in Table 1.

188

### 189 **SOPHiA DDM HRD Solution**

190 The HRD status was positive in 56 cases (64.4%), negative in 23 cases (26.4%), and  
191 undetermined in 8 cases (9.2%). The GI status was positive in 50 cases (57.5%), negative in  
192 27 cases (31.0%), and undetermined in 10 cases (11.5%). The BRCA status was positive in  
193 28 cases (32.2%), negative in 45 cases (51.7%), and undetermined in 14 cases (16.1%).

194

### 195 **OncoPrint Comprehensive Assay Plus Sequencing**

196 The mean of the average base coverage was 2469.64. The mean of the median absolute  
197 pairwise difference (MAPD) was 0.24. The MAPD is a metric that measures read coverage  
198 noise detected across all amplicons in a panel. A higher MAPD typically indicates lower  
199 coverage uniformity, which can result in missed or erroneous CNV calls. The quality control  
200 parameter metrics are summarized in Online Supplemental Table 1.

201

### 202 **BRCA1/2 Pathogenic Variants**

203 The concordance rate for BRCA1/2 pathogenic variants between the SOPHiA DDM HRD  
204 Solution and the OCA Plus was 95.9%. The discordant cases included two frameshift variants  
205 at homopolymer sequences. The OCA Plus pipeline filtered out these pathogenic variants due  
206 to an unusual prediction filter that measured the amount of strand bias according to the  
207 manufacturer's specifications. These two frameshift variants were restored by modifying the  
208 parameter of the unusual prediction filter. The other discordant variant was a long deletion.

209 The long deletion could not be detected because it spanned across the ends of amplicons. The  
210 pathogenic variants found in BRCA1/2 are listed in Online Supplemental Table 2.

211

## 212 **Selecting Model and Performance Estimation**

213 After excluding cases without a GI index from the SOPHiA DDM HRD Solution, the training  
214 set and test set consisted of 51 and 26 cases, respectively. The model with a penalty of 0.1  
215 and a mixture of 1 (Lasso regression) demonstrated the best performance in terms of the  
216 AUROC in a repeated fivefold cross-validation (Supplemental Figure 2). We fit the final  
217 model with the selected hyperparameters using the entire training set. The final model  
218 equation is  $-21.77 + 0.200 \times \text{GIM} + 0.102 \times \text{LOH}(\%) + 0.037 \times \text{TAI}(\%) + 0.261 \times \text{LST}(\%)$ .

219 To improve the performance of the prediction model, we added a borderline result category to  
220 the GI results. We classified cases with predicted values between -3 and 3 as borderline  
221 (Figure 1). The accuracy of our HRD status prediction model was 0.947 for the training set  
222 and 0.958 for the test set. Detailed performance metrics are summarized in Table 2.

223

## 224 **Genomic Instability Metric**

225 The AUROC for the GI status of SOPHiA DDM HRD Solution was 0.887 (Figure 2A). We  
226 set the positive cut-off value at 16 (Figure 2B). The accuracy of the HRD status using the  
227 GIM with a cut-off value of 16 was 0.911. Detailed performance metrics are summarized in  
228 Table 4.

229

## 230 **DISCUSSION**

### 231 **Summary of Main Results**

232 In this study, we developed a penalized linear regression model using the OCA Plus, which  
233 showed a high concordance rate with the SOPHiA DDM HRD Solution. We also observed  
234 that the GIM of the OCA Plus demonstrated high accuracy compared with the SOPHiA DDM  
235 HRD Solution. Despite being independently developed by different manufacturers, the GIM  
236 of the OCA Plus and the GI index of the SOPHiA DDM HRD Solution exhibited a high  
237 concordance rate. These findings suggest that both tests capture the same tumor characteristic,  
238 namely genomic alteration associated with HRD. When two different tests yield the same  
239 results, it reinforces the certainty of the results. It also indicates that both tests are reliable and  
240 reproducible.

241

#### 242 **Results in the Context of Published Literature**

243 It is important to note that neither the OCA Plus nor the SOPHiA DDM HRD Solution has  
244 been validated as a biomarker for PARP inhibitor response through clinical trials. Both the  
245 OCA Plus and the SOPHiA DDM HRD Solution require clinical validation through a clinical  
246 trial or a concordance test with a Food and Drug Administration–approved HRD test.

247 The SOPHiA DDM HRD Solution exhibited a considerable rate of failure, resulting  
248 in an undetermined result. This failure rate is similar to that of the myChoice HRD Plus  
249 assay<sup>15</sup>. Additionally, the OCA Plus fails to analyze HRD scores, and our prediction model  
250 relies on the OCA Plus HRD scores. Our prediction model includes a borderline category,  
251 which does not definitively determine the GI status. However, the GIM also demonstrated  
252 high accuracy without the need for a borderline category.

253

#### 254 **Strengths and Weaknesses**

255 In three cases (4%), the OCA Plus failed to detect BRCA1/2 pathogenic variants. Two of

256 these were c.2175del [chr13:32910667del] and c.3503dup [chr17:41244048dup] and were  
257 filtered out due to the application of a filter related to strand bias originating from  
258 homopolymer sequences. These false negatives were restored by modifying the filter  
259 parameter. The remaining one is the c.2593\_2621del [chr17:41244928\_41244956del]  
260 mutation, a 26 base pair deletion located within the overlapping regions of the OCA Plus  
261 amplicons. This type of long deletion seems to interfere with the generation of libraries  
262 containing both amplicons carrying this mutation, resulting in the absence of sequencing  
263 reads. It has been observed that the coverage depth of these two amplicons is relatively low  
264 compared with the adjacent amplicon positions. It is anticipated that detecting this mutation  
265 with the OCA Plus would be challenging. Therefore, interpretation of BRCA1/2 status results  
266 should consider the limitations of the test. These two cases had high GI and were classified as  
267 HRD positive.

268         Because the NGS study was not conducted on all patients with high-grade ovarian  
269 serous carcinoma, the patients included in this study may exhibit bias. However, the rates of  
270 BRCA1/2 pathogenic variant presence and positive HRD status are similar to those reported  
271 in a clinical trial of ovarian high-grade serous carcinoma using the myChoice HRD Plus  
272 assay (Myriad Genetic Laboratories).

273         We developed the penalized linear regression model by using a small training set and  
274 validated it with a small test set. This approach may lead to a model that is either too  
275 simplistic and underfits the data or too complex and overfits the data.

276

## 277 **Implications for Practice and Future Research**

278 The OCA Plus offers several advantages compared with HRD-specific tests. It enables  
279 comprehensive analysis of genetic alterations, including single nucleotide variants (SNVs),

280 insertions and deletions (indels), CNVs, structural variations, the tumor mutation burden,  
281 mismatch repair deficiency, and microsatellite instability. This broad coverage enhances the  
282 ability to identify potential targeted treatments. The high accuracy between the OCA Plus and  
283 the SOPHiA DDM HRD Solution supports its potential as a biomarker for predicting the  
284 PARP inhibitor response and its application in clinical trials for PARP inhibitors. Additionally,  
285 our study provides a cut-off value for the GIM of the OCA Plus that correlates with the  
286 SOPHiA DDM HRD Solution, with a high accuracy of 0.911.

287

## 288 **CONCLUSIONS**

289 This study presents a homologous recombination deficiency prediction model from the  
290 Oncomine Comprehensive Assay Plus that correlates with the SOPHiA DDM HRD Solution.  
291 this study provides evidences that The Oncomine Comprehensive Assay Plus provides  
292 reliable biomarkers for homologous recombination.

293

## 294 **REFERENCES**

- 295 1. Nguyen L, Martens J, Van Hoeck A, et al. Pan-cancer landscape of homologous  
296 recombination deficiency: *Cancer Biology*, 2020.
- 297 2. von Werdt A, Brandt L, Schärer OD, et al. PARP Inhibition in Prostate Cancer With  
298 Homologous Recombination Repair Alterations. *JCO Precision Oncology*  
299 2021(5):1639-49. doi: 10.1200/PO.21.00152
- 300 3. Scully R, Livingston DM. In search of the tumour-suppressor functions of BRCA1 and  
301 BRCA2. *Nature* 2000;408(6811):429-32. doi: 10.1038/35044000
- 302 4. Rahman N, Seal S, Thompson D, et al. PALB2, which encodes a BRCA2-interacting

- 303 protein, is a breast cancer susceptibility gene. *Nat Genet* 2007;39(2):165-67. doi:  
304 10.1038/ng1959
- 305 5. Sharan SK, Morimatsu M, Albrecht U, et al. Embryonic lethality and radiation  
306 hypersensitivity mediated by Rad51 in mice lacking Brca2. *Nature*  
307 1997;386(6627):804-10. doi: 10.1038/386804a0
- 308 6. Hussain M, Mateo J, Fizazi K, et al. Survival with Olaparib in Metastatic Castration-  
309 Resistant Prostate Cancer. *New England Journal of Medicine* 2020;383(24):2345-57.  
310 doi: 10.1056/NEJMoa2022485
- 311 7. Litton JK, Rugo HS, Ettl J, et al. Talazoparib in Patients with Advanced Breast Cancer and  
312 a Germline BRCA Mutation. *New England Journal of Medicine* 2018;379(8):753-63.  
313 doi: 10.1056/NEJMoa1802905
- 314 8. Moore K, Colombo N, Scambia G, et al. Maintenance Olaparib in Patients with Newly  
315 Diagnosed Advanced Ovarian Cancer. *New England Journal of Medicine*  
316 2018;379(26):2495-505. doi: 10.1056/NEJMoa1810858
- 317 9. Robson M, Im S-A, Senkus E, et al. Olaparib for Metastatic Breast Cancer in Patients with  
318 a Germline BRCA Mutation. *New England Journal of Medicine* 2017;377(6):523-33.  
319 doi: 10.1056/NEJMoa1706450
- 320 10. Abkevich V, Timms KM, Hennessy BT, et al. Patterns of genomic loss of heterozygosity  
321 predict homologous recombination repair defects in epithelial ovarian cancer. *British*  
322 *Journal of Cancer* 2012;107(10):1776-82. doi: 10.1038/bjc.2012.451
- 323 11. Birkbak NJ, Wang ZC, Kim J-Y, et al. Telomeric Allelic Imbalance Indicates Defective  
324 DNA Repair and Sensitivity to DNA-Damaging Agents. *Cancer Discov*  
325 2012;2(4):366-75. doi: 10.1158/2159-8290.CD-11-0206
- 326 12. Popova T, Manié E, Rieunier G, et al. Ploidy and Large-Scale Genomic Instability

- 327 Consistently Identify Basal-like Breast Carcinomas with BRCA1/2 Inactivation.  
328 *Cancer Res* 2012;72(21):5454-62. doi: 10.1158/0008-5472.CAN-12-1470
- 329 13. Telli ML, Timms KM, Reid J, et al. Homologous Recombination Deficiency (HRD)  
330 Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in  
331 Patients with Triple-Negative Breast Cancer. *Clin Cancer Res* 2016;22(15):3764-73.  
332 doi: 10.1158/1078-0432.CCR-15-2477
- 333 14. Watkins JA, Irshad S, Grigoriadis A, et al. Genomic scars as biomarkers of homologous  
334 recombination deficiency and drug response in breast and ovarian cancers. *Breast*  
335 *Cancer Research* 2014;16(3):211. doi: 10.1186/bcr3670
- 336 15. Ray-Coquard I, Pautier P, Pignata S, et al. Olaparib plus Bevacizumab as First-Line  
337 Maintenance in Ovarian Cancer. *New England Journal of Medicine*  
338 2019;381(25):2416-28. doi: 10.1056/NEJMoa1911361
- 339 16. de Bono J, Mateo J, Fizazi K, et al. Olaparib for Metastatic Castration-Resistant Prostate  
340 Cancer. *New England Journal of Medicine* 2020;382(22):2091-102. doi:  
341 10.1056/NEJMoa1911440
- 342 17. Andre G, Coletta T, Pozzorini C, et al. A deep learning approach for improved detection  
343 of homologous recombination deficiency from shallow genomic profiles: bioRxiv,  
344 2022.
- 345
- 346



347 Table 1. Cases summary

	Overall	Training set	Test set	p value
Number of patients	87	55	32	
Age (mean (SD))	59.3 (10.3)	59.2 (10.5)	59.4 (10.2)	0.946
FIGO stage (%)				0.875
1	7 (8.1)	5 (9.3)	2 (6.2)	
2	6 (7.0)	3 (5.6)	3 (9.4)	
3	63 (73.3)	40 (74.1)	23 (71.9)	
4	10 (11.6)	6 (11.1)	4 (12.5)	
Institution (%)				
Hospital A	55 (63.2)	55 (100.0)	0 (0.0)	
Hospital B	15 (17.2)	0 (0.0)	15 (46.9)	
Hospital C	10 (11.5)	0 (0.0)	10 (31.2)	
Hospital D	4 (4.6)	0 (0.0)	4 (12.5)	
Hospital E	3 (3.4)	0 (0.0)	3 (9.4)	

348 SD: standard deviation, FIGO: The International Federation of Gynecology and Obstetrics

349

350 Table 2. Performance metrics

HRD test	N	Accuracy	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	F1 score
HRD result (Train)	38	0.947 95% CI ( 0.823 - 0.994 )	0.964	0.900	0.964	0.900	0.964
HRD result (Test)	24	0.958 95% CI ( 0.789 - 0.999 )	0.938	1.000	1.000	0.889	0.968
HRD result (GIM)	79	0.911 95% CI ( 0.826 - 0.964 )	0.964	0.783	0.915	0.900	0.939
GI result (Train)	33	0.97 95% CI ( 0.842 - 0.999 )	1.000	0.909	0.957	1.000	0.978
GI result (Test)	23	0.87 95% CI ( 0.664 - 0.972 )	0.786	1.000	1.000	0.750	0.880
GI result (GIM)	77	0.883 95% CI ( 0.79 - 0.945 )	0.960	0.741	0.873	0.909	0.914

351 HRD: homologous recombination deficiency, HRD result (Train/Test): prediction model performance for HRD result of SOPHiA DDM

352 HRD Solution, GI: Genomic Integrity, GI Result (Train/Test): prediction model performance for GI result of SOPHiA DDM HRD Solution

353 F1 score: harmonic mean of the precision and recall, GIM: Genomic Instability Metric, HRD result/GI result (GIM): prediction performance

354 of GIM for HRD result/GI result of SOPHiA DDM HRD Solution

355 **Acknowledgement**

356 The authors thank Taeun Kim for providing data acquisition.

357

358 **Declaration of Conflicting Interests**

359 The Authors declares that there is no conflict of interest.

360

361 Figure legends

362 Figure 1. Performance of the homologous recombination deficiency (HRD) prediction model.

363 The training set (A) and the test set (B) are shown. The black vertical lines represent

364 borderline cut-off values (-3 and 3).

365 Figure 1. Receiver operating characteristic curve (ROC) for the genomic instability (GI)

366 status of the SOPHiA DDM HRD Solution (A). The threshold (cut-off value) was set at 16.

367 The black vertical line indicates a genomic instability metric (GIM) of 16. The black

368 horizontal line indicates a GI index of 0.



