Performance of an Optimized Methylation-Protein Multi-Cancer Early Detection (MCED) Test Classifier

Vladimir G. Gainullin,[1] Melissa Gray, [1] Madhav Kumar,[1] Stephen Luebker,[1] Amy M. Lehman,[1] Omair A. Choudhry,[1] Julia Roberta,[1] Darl D. Flake II,[1] Avinash Shanmugam,[1] Kevin Cortes,[1] Emily Chang,[1*] Philip J. Uren,[2] Amin Mazloom,[2] Jorge Garces,[2] Gerard A. Silvestri,[3] David W. Chesla,[4] Robert W. Given,[5] Tomasz M. Beer,[1] Frank Diehl[2**]

[1]Exact Sciences Corporation, Madison, WI, USA
[2]Exact Sciences Corporation, La Jolla, CA, USA
[3]Medical University of South Carolina, Charleston, SC, USA
[4]Corewell Health, Grand Rapids, MI, USA
[5]Urology of Virginia, Virginia Beach, VA, USA
*Former exact Sciences employee

**Corresponding Author
Frank Diehl, PhD
Senior Vice President, Research & Development, Multi-cancer Early Detection
Exact Sciences Corporation
11085 N. Torrey Pines Rd.
La Jolla, CA 92037
Email: fdiehl@exactsciences.com

Short title Multi-cancer early detection biomarkers in plasma

Keywords Multi-cancer early detection, circulating tumor DNA (ctDNA), methylation, proteins, machine learning

Abstract

Multi-cancer early detection (MCED) tests can detect several cancer types and stages. We previously developed a methylation and protein (MP V1) MCED classifier. In this study, we present a refined MP V2 classifier, developed by evaluating model architectures that improved performance in prospectively enrolled case-control cohorts under standard testing conditions. The newly developed MP V2 classifier was trained to be more generalizable and achieve increased early-stage sensitivity at a target specificity of ≥97.0%. MP V1 and MP V2 classifier performances were compared using a previously described test set, and MP V2 performance was also evaluated in a new independent clinical validation set. Compared to MP V1, the MP V2

classifier demonstrated a 7.3% increase in overall sensitivity, with sensitivity increases of 7.6%, 9.2%, and 8.3% for stages I, II, and stages I/II, respectively, in the intended use (breast and prostate cancers excluded) test set. In an independent validation intended use set, the MP V2 classifier showed an overall sensitivity of 55.6%, with sensitivities of 26.8%, 42.9%, and 34.8% for stages I, II, and stages I/II, respectively. In a case-control setting, the MP V2 classifier offered improved sensitivity for early-stage cancers at a lower specificity target.

Introduction

Cancer is the second leading cause of death in the United States, with more than 2 million new cancer cases and over 626,000 deaths projected for 2026.[1] Guideline-recommended single-cancer screening is available for only four cancers: breast, cervical, colorectal, and lung.[2-5] Such screening has resulted in early cancer detection and contributed to improved clinical outcomes and mortality for these cancer types.[1] However, guideline-recommended single-cancer screening addresses only about 1/3 of incident cancers and cancer related deaths,[6] with 2/3 lacking recommended screening options.[1] Finally, current single-cancer screening tests sacrifice specificity to optimize sensitivity, resulting in substantial false positive (FP) rates. The reported FP rates for guideline-recommended single-cancer screening tests range from 6.0% for colorectal (age-adjusted mt-sDNA) to 14.5% for cervical (Papanicolaou test) for each round of screening.[7-10] Since the probability of obtaining a FP result increases with each round of recommended breast, cervical, colorectal, and lung cancer screening, the cumulative estimated lifetime FP risk associated with guideline-recommended single-cancer screening is considerable.[11, 12] To address some of these limitations and improve upon current population-based single-cancer screening efforts, significant efforts are being made in the development of early-detection screening strategies that simultaneously detect multiple cancers in a single blood test.

66 Advances in sensitive high-throughput technologies and machine learning have led to the
67 development of minimally invasive, blood-based, single and multi-biomarker class multicancer
68 early detection (MCED) tests. Several MCED tests are under development to identify multiple
69 cancers earlier by detecting shared cancer-derived DNA (methylation, fragmentation,
70 mutations), proteins, or other biomarkers in the blood.[13-19] Since MCED tests interrogate cancer
71 signals originating from multiple organ systems they represent a potential paradigm shift in
72 cancer screening, and several assay attributes must be considered to optimize test
73 development.
74
75 First, MCED tests must expand the range of screen detected cancers. Second, MCED tests
76 must demonstrate sensitivity to detect early-stage cancers, particularly cancers where early-
77 stage diagnoses enable intervention with curative intent. Third, sensitivity must be balanced with
78 a relatively high specificity. For MCED tests, this balance aims to maximize potential benefits
79 (i.e., early-stage detection) while keeping harms from FP results acceptable. These harms may
80 be further reduced by an efficient diagnostic strategy following a positive MCED test that
81 minimizes the diagnostic burden for patients.
82
83 Sensitivity for early-stage solid tumors, when treatment with curative intent is available, is likely
84 to have an outsized impact on the benefits of screening with MCED. Notably, in the Detecting
85 cancers Earlier Through Elective mutation-based blood Collection and Testing (DETECT-A)
86 trial, the first prospective, interventional study to investigate MCED testing, all eight patients
87 whose cancers were treated following detection via MCED testing and subsequent diagnosis at
88 Stage I or II remain alive and cancer-free after >4 years of follow up.[20] While of modest size, this
89 observation highlights the feasibility of early detection of solid tumors using a multi-biomarker
90 class MCED test, and provides early evidence that successful treatment can be reliably

91 delivered. These observations highlight the importance of an emphasis on early-stage sensitivity
92 in MCED development. [19]
93 We previously developed a methylation and protein (MP V1) MCED classifier utilizing a
94 prospectively collected case-control study of individuals aged 50-84, allowing the detection of a
95 broad range of cancer organ types at all stages.[15] The specificity target for MP V1 was set at a
96 minimum of ≥98.0%. Subsequent validation using an independent test set confirmed that the
97 MP V1 classifier achieved the specificity requirement, with observed sensitivities of 15.4%,
98 38.0%, and 26.1% for stage I, stage II, and stage I and II combined, respectively.[15]
99
100 Molecular diagnostic tests, including MCED tests, are often developed using case-control
101 studies.[21, 22] MCED tests are designed with high specificity targets to limit FP results, and the
102 performance observed using case-control samples may not accurately represent real-world
103 performance in a clinical setting.[17, 18, 21, 22] To demonstrate clinical utility, MCED tests must be
104 designed to detect cancers at early stages when they may be more amenable to curative intent.
105 In this algorithm development study, we developed an optimized MP classifier architecture (MP
106 V2) using a lower target specificity (≥97.0%) in a case-control setting to optimize sensitivity for
107 early-stage cancers. We then validated the classifier performance using an independent testing
108 set that more closely resembled the intended-use population.

109 **Results**

110 *Classifier Selection using an Independent Mini-holdout (MHO) Test Set*

111 MP V1 and MP V2 performance were first examined in MHO classifier selection test set
112 samples from a previously untested cohort of Ascertaining Serial Cancer patients to Enable
113 New Diagnostic 2 (ASCEND-2) study subjects (Figure 1). A series of comparisons were
114 performed to examine changes in target specificity, classifier model architecture, and both
115 together.

Lowering the specificity target and using the MP V1 classifier

Performance is shown in Table 1 and Supplementary Materials Table S1. At a target specificity of ≥98%, the MP V1 classifier resulted in a measured specificity of 99.4% (95% confidence interval [95%CI]: 98.3, 99.8), with an overall sensitivity of 60.9% (95% CI: 51.6, 69.5), and sensitivities of 31.8% (95% CI:16.4, 52.7), 47.7% (95% CI: 33.8, 62.1), 63.6% (95% CI: 43.0, 80.3), 68.2% (95% CI: 47.3, 83.6), 95.7% (95% CI: 79.0, 99.2), and 42.9% (95% CI: 24.5, 63.5) for stages I, I and II combined, II, III, IV, and unknown, respectively for the MHO classifier selection test set (Table 1). When the MP V1 target specificity was lowered to ≥97.0%, the MP V1 classifier (measured specificity of 99.0%; 95% CI: 97.7, 99.6) demonstrated an overall sensitivity of 64.5% (95% CI: 55.3, 72.9), and sensitivities of 36.4% (95% CI: 19.7, 57.0), 52.3% (95% CI: 37.9, 66.2), 68.2% (95% CI: 47.3, 83.6), 77.3% (95% CI: 56.6, 89.9), 95.7% (95% CI: 79.0, 99.2), and 42.9% (95% CI: 24.5, 63.5) for stages I, I and II combined, II, III, IV, and unknown, respectively for the MHO classifier selection test set (Supplementary Materials Table S1).

By lowering the MP V1 target specificity by 1% (≥98.0% to ≥97.0%), we observed absolute sensitivity increases of 3.6%, 4.6%, 4.6%, and 4.6% for overall, stage I, II, and stages I and II combined, respectively, in the MHO classifier selection test set (Table 1). These sensitivity increases are equivalent to relative sensitivity increases of 5.9%, 14.5%, 7.2%, and 9.6% for overall, stage I, II, and stages I and II combined, respectively.

Incorporating improved classifier model architecture and a lower specificity target

The MP V2 classifier at a target specificity of ≥97.0% (measured specificity of 99.0%; 95% CI: 97.7, 99.6) demonstrated an overall sensitivity of 72.7% (95% CI: 63.7, 80.2), and sensitivities of 40.9% (95% CI: 23.3, 61.3), 56.8% (95% CI: 42.2, 70.3), 72.7% (95% CI: 51.8, 86.8), 86.4% (95% CI: 66.7, 95.3), 100.0% (95% CI: 85.7, 100.0), and 61.9% (95% CI: 40.9, 79.2) for stages

142 I, I and II combined, II, III, IV, and unknown, respectively in the MHO classifier selection test set
143 (Supplementary Materials Table S1).
144
145 *Evaluating the impact of changes to the specificity target and changes to the model architecture*
146 MP V2 classifier performance (target specificity of ≥97%) was compared to the MP V1 classifier
147 performance (target specificity of ≥98%). Compared to the MP V1 classifier (measured
148 specificity of 99.4%), the MP V2 classifier (measured specificity of 99.0%) resulted in absolute
149 sensitivity increases of 11.8%, 9.1%, 9.1%, and 9.1% for overall, stage I, II, and stages I and II
150 combined, respectively, in the MHO classifier selection test set (Table 1). These absolute
151 sensitivity increases are equivalent to relative sensitivity increases of 19.4%, 28.6%, 14.3%, and
152 19.1% for overall, stage I, II, and stages I and II combined, respectively (Table 1),
153
154 *Isolating the impact of model architecture evolution*
155 To evaluate the performance contributions that may be attributed to improved MP V2 model
156 architecture versus a lowered specificity, MP V2 and MP V1 classifier performances were
157 compared at the same target specificity of ≥97% (measured specificity of 99.0%). Compared to
158 MP V1, MP V2 demonstrated absolute sensitivity increases of 8.2%, 4.5%, 4.5%, and 4.5% for
159 overall, stage I, II, and stages I and II combined, respectively, in the MHO classifier selection
160 test set. These absolute sensitivity increases are equivalent to relative sensitivity increases of
161 12.7%, 12.4%, 6.6%, and 8.6% for overall, stage I, II, and stages I and II combined, respectively
162 (Table 1). The performance results above suggest that a substantial portion of the observed
163 increases in stage I/II sensitivity may be attributed to MP V2 model architecture improvements.
164
165 *MP V1 and V2 Classifier Test Set Performance Comparison*
166 The locked MP V2 classifier was then evaluated using the original ASCEND-2 test development
167 set cohort that was used to evaluate the previously reported MP V1 classifier performance

168 (Figure 1).[15] At a measured specificity of 97.4%, MP V2 demonstrated improvements across all
169 stage groups compared to MP V1 (specificity 98.5%) (Table 2). Overall sensitivity increased
170 from 50.9% (MP V1) to 57.8% (MP V2), with stage-specific sensitivity increases of 6.0% for
171 stage I, 8.0% for stage II, and 6.9% for stage I–II combined. Relative sensitivity increases were
172 39.0% for stage I, 21.1% for stage II, and 26.4% for stage I–II combined. Detailed results by
173 stage are shown in Table 2. The differences in absolute sensitivity were statistically significant
174 by McNemar's test ($p < 0.01$) for all comparisons except stage IV ($p=0.039$) and stage unknown
175 ($p=0.125$). V2 and V1 methylation sub-models demonstrated similar discrimination (AUROC
176 0.83 and 0.81, respectively) for this dataset. At equivalent targeted specificity (>97.0%), MP V2
177 classifier showed improvements compared to MP V1 in early-stage sensitivity (33.0% vs 31.0%)
178 and specificity (97.4% vs 97.2%), however these results were not statistically significant.
179
180 MP V1 and MP V2 classifier performance was also compared using the test development set
181 with breast and prostate cancers excluded. At a specificity of 97.4%, MP V2 demonstrated
182 improvements across all stage groups compared to MP V1 (specificity 98.5%) (Table 2). Overall
183 sensitivity increased from 56.8% (MP V1) to 64.1% (MP V2), with stage-specific improvements
184 of 7.6% for stage I, 9.2% for stage II, and 8.3% for stage I–II combined. Relative sensitivity
185 increases were 44.2% for stage I, 18.9% for stage II, and 27.0% for stage I–II combined.
186 Detailed results by stage are shown in Table 2. The differences in absolute sensitivity were
187 statistically significant by McNemar's test ($p < 0.01$) for all comparisons except stage IV
188 ($p=0.070$) and stage unknown ($p=0.125$).
189
190 Overall sensitivities by organ type for MP V1 (measured 98.5% specificity) and MP V2
191 (measured 97.4% specificity) classifiers using the test development set ranged from 11.8% for
192 prostate to 80.0% for liver and bile duct (Supplementary Materials Figure S2). MP V1 and MP
193 V2 performance comparisons by organ type and stage are shown in Supplementary Materials

194 Table S2. MP V2 absolute increase in sensitivity was statistically significant for breast (p-
195 value=0.016), bladder (p-value=0.031), and colorectal (p-value=0.002) cancers.
196
197 *MP V2 Independent Clinical Validation Set: Participant Characteristics*
198 MP V2 classifier performance was also evaluated in an independent clinical validation
199 (N=1,124; 324 cancer and 800 non-cancer) test set consisting of previously untested samples
200 from the ASCEND-2 study (Figure 1). The clinical validation set consisted of 1,124 subjects with
201 mean age of 64.8 years; 51.2% were female. 83.7% of the clinical validation cohort identified as
202 White, 10.8% Black or African American, and 1.1% Asian; 83.6% were of non-Hispanic/Latino
203 ethnicity. Fifty-three percent were never cigarette smokers, and 50.4% indicated current alcohol
204 consumption. The mean body mass index was 30.0 kg/m^2, and 56.3% indicated a family history
205 of cancer in a first-degree relative. Cancer and non-cancer participants had similar age, sex,
206 and race/ethnicity, body mass index, and cigarette smoking/alcohol consumption distributions
207 (Table 3). This sample set was selected to more closely resemble the intended use population.
208
209 *MP V2 Classifier Independent Clinical Validation Set Performance*
210 Using the independent clinical validation set, the MP V2 classifier, at a measured specificity of
211 97.4% (95% CI: 96.0, 98.3), showed an overall sensitivity of 41.4% (95% CI: 36.1, 46.8), and
212 sensitivities of 16.0% (95% CI: 10.3, 24.2), 22.8% (95% CI: 17.4, 29.2), 31.3% (95% CI: 22.4,
213 41.9), 52.9% (95% CI: 41.3, 64.1), and 83.1% (95% CI: 72.2, 90.3) for stages I, I and II
214 combined, II, III, and IV, respectively (Table 4). Overall sensitivity excluding cancer organ types
215 with average-risk, single-cancer screening (i.e. excluding breast, prostate, cervix, colon and
216 rectum) was 52.4% (95% CI: 43.7, 60.9) and 67.9% (95% CI: 58.5, 76.0) for the six most
217 aggressive cancer organ types with the shortest 5-year survival rate (i.e. pancreas, esophagus,
218 liver, lung and bronchus, stomach, and ovary).[23] With breast and prostate cancers excluded, the
219 MP V2 classifier demonstrated an overall sensitivity of 55.6% (95% CI: 49.0, 62.0), and

220 sensitivities of 26.8% (95% CI: 17.0, 39.6), 34.8% (95% CI: 26.6, 44.0), 42.9% (95% CI: 30.8,
221 55.9), 63.6% (95% CI: 50.4, 75.1), and 89.3% (95% CI: 78.5, 95.0) for stages I, I and II
222 combined, II, III, and IV, respectively (Table 4). The MP V2 classifier overall sensitivities by
223 organ type ranged from 3.9% (prostate) to 100.0% (ovary, small intestine, anus, vulva, and
224 testis) (Supplementary Materials Table S4) in the clinical validation set. Table S4 also shows the
225 performance by stage.
226
227 *Cancer Subtype Analysis*
228 The MP test is based on sets of shared methylation and protein biomarkers present in many
229 different tumor types; these biomarkers have the potential to detect a broad range of cancers
230 from a blood specimen. The case report forms obtained from 635 subjects with cancer (421
231 from the MP V2 test development set, 80 from the MHO test set, and 134 from the independent
232 clinical validation test set) were assessed to determine the tumor types and histological
233 subtypes detected. Over fifty distinct cancer subtypes were identified. Among those, 17 cancer
234 subtypes were identified based on single cases (n=1). Cancer subtype staging was not
235 determined. The cancer subtypes detected are listed in Supplementary Materials Table S5.
236
237 **Discussion**
238 MCED testing has the potential to shift population-wide cancer screening from single-cancer
239 screening of a handful of cancers to many cancers through expanded multi-cancer screening
240 using a single blood specimen. MCED tests are typically designed with high specificity targets to
241 limit FP results and the associated harms associated with futile diagnostic procedures.[24]
242 Classifier development is challenging as it requires balancing case-control derived
243 sensitivity/specificity measurements to project real world-clinical performance. We therefore
244 attempted to develop a refined classifier that is more generalizable to cohort differences and
245 target a lower target specificity in a case-control setting. In this study, we first identified and

246 selected a new model architecture and then adjusted the specificity target to evaluate
247 performance in a classifier selection set, a previously described test set, and finally in an
248 independent validation set more closely representative of the intended use population. We
249 found that incorporating MP V2 model architecture improvements and lowering the specificity
250 target by one percent resulted in a 6.9% absolute (26.4% relative) sensitivity increase for early-
251 stage cancers (stage I and II combined) compared to the MP V1 classifier. A detailed analysis of
252 the MHO classifier selection set showed that the observed performance improvements are due
253 to both the lower case-control target specificity and the new model architecture. The MP V2
254 classifier will be utilized in a real-world setting to further validate the sensitivity and specificity
255 results.
256
257 False positive results are of concern for both single cancer and MCED tests. However, since
258 MCED tests interrogate cancer signals originating from multiple organ sites, FP MCED results
259 could trigger diagnostic procedures that may impose physical and psychological burdens on
260 patients to definitively rule in or rule out cancer. Specificities of ≥98.9% have been reported for
261 the prospective PATHFINDER, PATHFINDER-2, and DETECT-A studies, resulting in FP rates
262 of ≤2%,[17, 18, 19] much lower than FP rates reported for guideline-recommended single-cancer
263 screening tests.[12] An invasive procedure (e.g., surgery) was required to resolve FP results in
264 just 2%, 0.4%, and 3% of cases, respectively in the aforementioned studies.[17, 18, 19] Efficient
265 imaging-based diagnostic resolution strategies may further reduce the potential risks associated
266 with the workup for FP MCED test results.[25]
267
268 This study has some limitations. Of the projected two million new cancer diagnoses in 2026,[1] it
269 is estimated that approximately 88% (almost 1.8 million) will occur in individuals aged 50 years
270 and older.[26] In addition, it is likely that the prevalence of undiagnosed cancers harbored by
271 apparently healthy individuals at any point in time is higher as many cancers have pre-clinical

272 dwell times that extend beyond one year.[27] The controls used to develop and test the MP V1
273 classifier included age-matched individuals without suspicion of cancer. While free from
274 suspicion of cancer at the time of enrollment, controls were not required to have completed all
275 recommended cancer screening tests, were not subject to any procedures to rule out the
276 presence of an occult cancer and participants were not followed after contributing a blood
277 sample to confirm the absence of cancer that might emerge in subsequent months. Therefore,
278 the control group likely included individuals with hidden, undiagnosed cancers. Thus, targeting a
279 specificity of ≥98.5% in the MP V1 classifier development case-control study may have been
280 conservative, given that measured specificity was likely slightly underestimated by undiagnosed
281 cancers in the control group, potentially sacrificing sensitivity.
282
283 Although the ASCEND-2 samples used in this study were prospectively collected from more
284 than 200 sites across the US and Europe and included a diverse participant cohort broadly
285 representative of the U.S. population, the analyses were performed with cases having tumor
286 type representation that is close to an intended use population but with equal stage distributions
287 across several cancer types. Therefore, the ASCEND -2 case-control performance results might
288 be different than real-world test performance. Both the MP V1 and MP V2 classifiers were
289 evaluated using the same test set design. When MP V2 classifier sensitivity was assessed in a
290 clinical validation cohort that more closely reflected SEER incidence, the observed overall
291 sensitivity was 41.4%. MCED test performance must be interpreted with considerable caution,
292 as differences in cancer types/stages affect performance estimates. Previous studies have
293 reported significant differences in performance between case-control and prospective
294 interventional studies,[17-19, 21, 22] even for the same assay in different prospective studies.[18, 19]
295 While improvements in performance due to test development can be assessed in case-control
296 studies, as reported here, performance in the screening setting is determined in prospective
297 studies in the intended-use population. Therefore, performance comparisons across studies

298 should account for differences in design, cancer type mix, and stage distribution, and simple
299 cross-study comparisons should be avoided.
300
301 In summary, as part of an ongoing MCED test development process, we trained and tested a
302 refined MCED classifier, MP V2. We adjusted a pre-specified case-control specificity target from
303 ≥98% to ≥97% to optimize early-stage (stages I and II) cancer detection which should result in
304 maintained or increased specificity in a real-world setting.[17, 18, 21, 22] The locked MP V2 classifier
305 was then evaluated in a clinical validation cohort selected to more closely resemble a real-world
306 population. Prospective, real-world evaluation of the MP V2 classifier will be needed to
307 determine if we achieved the optimal balance between sensitivity and specificity.
308
309 **Methods**
310 *MP V1 Classifier*
311 A prospectively collected, case-control study, ASCEND-2, was the primary sample source for
312 the development of the MP V1 classifier.[15] MP V1 classifier development included participants
313 from multiple sites across the United States and within Europe. The MP V1 classifier
314 development cohort included male and female subjects ≥50 years old with known cancer and
315 controls without suspicion of cancer. Cases were recruited immediately following cancer
316 diagnosis and before treatment initiation or, in the case of cancer types that are diagnosed with
317 excisional biopsies, immediately prior to the diagnostic and therapeutic procedure. All collection
318 sites had Institutional Review Board approval, and all eligible subjects provided written informed
319 consent and were assessed for study participation eligibility. Blood samples were collected in
320 LBgard® tubes at the study sites, shipped to centralized laboratories, and processed into plasma
321 and buffy aliquots. Protein and methylation assays were assessed using separate plasma
322 aliquots.
323

324 *Protein Analysis*

325 Plasma protein concentrations were measured using commercially available assays for 6
326 proteins on a Roche cobas e801 instrument (Roche Diagnostics, Indianapolis, IN) per
327 manufacturer's recommendations.[28] The cobas e801 analyzer uses electrochemiluminescence
328 (ECL)-based immunoassays to measure protein concentrations.

329

330 *Methylated DNA Marker (MDM) Analysis*

331 Methylated DNA markers were measured as previously described.[28] Briefly, circulating tumor
332 (ctDNA) was extracted from 6 mL plasma followed by bisulfite conversion on a Hamilton
333 Microlab STARlet (Hamilton Robotics, Reno, NV) using reagents manufactured by Exact
334 Sciences Corporation (Madison, WI). Bisulfite converted ctDNA was then analyzed using Target
335 Enrichment Long-probe Quantitative Amplified Signal (TELQAS™) chemistry. For the initial
336 target enrichment, converted ctDNA was transferred to a 96-well polymerase chain reaction
337 (PCR) plate and subjected to multiplex PCR to amplify the MDMs and a reference methylation
338 marker. The PCR products were diluted 10-fold and evaluated in triplex or biplex assays using
339 Long-probe Quantitative Amplified Signal (LQAS™) on a Quantstudio 5 Dx Real-Time PCR
340 Instrument (Life Technologies Corporation, Grand Island, NY). TELQAS combines amplification
341 using real-time PCR and allele-specific detection of methylated target DNA through an invasive
342 cleavage assay.

343

344 *MP V2 Classifier Training*

345 The overall MP V2 classifier development workflow is shown in Figure 1. The MP V2 classifier
346 candidates were trained using the same methylation and protein data set previously used for the
347 training of MP V1[15]. The total training set was comprised of 3,027 samples (654 cancer and
348 2,373 non-cancer), that was partitioned into two sub-groups for 5-fold cross validation (N=1,819;
349 431 cancer and 1,388 non-cancer) and MHO testing (N=1,208; 223 cancer and 985 non-

350 cancer). The MP V2 training set was used to investigate multiple alternative classifier
351 architectures, feature engineering, and transformation approaches as well as to perform tuning
352 of specificity thresholds. A target specificity of ≥97.0% was used for MP V2 classifier training.
353 Only samples that passed pre-specified methylation and protein assay quality control criteria
354 were included in the training process; these criteria were identical for the MP V1 and MP V2
355 training processes. Methylation and protein classifiers were trained individually, and the
356 individual classifier results were combined using a logical OR overarching classifier to generate
357 the final (positive or negative) call. The result was deemed positive when either the methylation
358 or protein assay provided a cancer signal detected call.
359
360 *MP V2 Classifier Model Selection*
361 After multiple potential classifiers were developed and evaluated as part of training, locked
362 candidate models were compared during a classifier selection step using newly generated
363 methylation and protein data from an independent MHO classifier selection test set (MHO;
364 n=619; n=110 Cancers; n=509 non-cancers). The samples were selected from a previously
365 untested cohort of subjects of the ASCEND-2 study cohort.[15] The performance of the MP V2
366 candidate classifiers was compared to the MP V1 classifier at two different target specificities
367 (original target of ≥98.0% and updated target of ≥97.0%). A Feature Weighted Naive Bayes
368 (FWNB) classifier was selected as it showed improved performance compared to both MP V1
369 and other candidate MP V2 model architectures. This model was locked for additional tuning
370 and then subjected to additional testing.
371
372 *MP V2 Classifier Evaluation*
373 To enable direct comparison of V2 to V1 MP classifiers, the locked MP V2 classifier was tested
374 using the original data set used for the performance characterization of the MP V1 classifier.[15]
375 This test set included (n=3,163; 729 cancers and 2,434 non-cancers) and was described above.

376 The MP V2 performance assessment at the adjusted specificity target (≥97%) was compared to
377 the MP V1 classifier at the original (≥98%) specificity target.
378
379 *MP V2 Classifier Clinical Validation*
380 The MP V2 classifier performance was further evaluated in an independent clinical validation
381 (N=1,124; 324 cancer and 800 non-cancer) test set. The samples for this set were selected
382 from previously untested subjects of the ASCEND-2 study cohort.[15] The clinical validation test
383 set cancer cohort included nineteen organ sites (prostate (n=51), breast (n=50), lung and
384 bronchus (n=57), colon and rectal (n=38), kidney (n=19), uterus (n=17), head and neck
385 (squamous cell; n=17), pancreas (n=17), bladder and urinary (n=12), liver (n=12), thyroid (n=5),
386 stomach (n=7), ovary (n=7), esophagus (n=6), small intestine (n=3), cervix uteri (n=2), anus
387 (n=2), vulva (n=1) and testis (n=1)). The stage distributions for these organ tumor sites are
388 shown in Supplementary Materials Table S3.
389
390 Cancer samples were selected to achieve balanced representation across stages I-IV
391 (excluding breast and prostate cancer) for estimation of stage-specific performance. Within
392 stage, the distribution of seventeen organ sites (excluding breast and prostate) was chosen to
393 reflect 2021 incidence reported by the SEER database.[23] For breast and prostate cancer, we
394 selected fifty samples each, ensuring the stage distribution reflected 2021 SEER incidence.
395 Non-cancer samples were chosen so that the distribution of age groups, sex, and race mirrored
396 the demographics reported in the 2021 Census, to more closely represent the intended use
397 population in a real-world setting.[29] The clinical validation participant demographics and clinical
398 characteristics are shown in Table 3.
399
400 *Statistical Analysis*

Sensitivity and specificity were estimated overall and stratified by cancer stage. Sensitivity was defined as the proportion of participants with disease who tested positive, and specificity as the proportion of participants without disease who tested negative. Two-sided 95% confidence intervals were calculated using the Wilson score method. Comparisons of sensitivity and specificity between classifiers were performed using McNemar's test for paired binary outcomes, applied within each stage group and for the overall cohort. Statistical significance was assessed using two-sided tests with a threshold of $p < 0.05$. SAS software (https://www.sas.com) was used for CV analysis, Python SciPy and statsmodels (https://www.statsmodels.org), scikit-learn (https://www.scikit-learn.org) was used for model development.

Cancer Subtype Analysis

Case report forms were assessed by three independent physicians to identify biopsy and histopathology variables for each detected tumor type across the cohorts. The cancers were classified into distinct histological types and subtypes based on cancer site, morphology, and international classification of diseases for oncology (ICD-O) codes. Certain cancer types were designated as unique due to differences in histology, pathogenesis, risk factors, and clinical management, reflecting their recognition as separate entities in clinical practice. Limited cancer type, histology subtype, and clinical information were available for this analysis, and certain histological subtypes were not classified as distinct cancer types. With limited numbers of cases that represent many cancer subtypes, these data should not be used to estimate real world test performance by cancer subtypes.

Data Availability

The data generated in this study is available upon request from the corresponding author.

Acknowledgements

The authors would like to thank Michelle Beidelschies for leadership and feedback. Carolyn Hall and Feyza Sancar (Exact Sciences Corporation, Madison, WI) provided manuscript writing and editorial support. The authors would also like to thank the patients who generously participated in the original sample collection studies and the principal investigators and institutions who oversaw their enrollment. This study was funded by Exact Sciences Corporation.

Author Contributions

Conceptualization: Frank Diehl

Methodology: Vladimir Gainullin, Frank Diehl

Formal Analysis: Vladimir Gainullin, Melissa Gray, Madhav Kumar, Stephen Luebker, Amy Lehman, Omair Choudhry, Julia Roberta, Darl Flake, Avinash Shanmugam, Kevin Cortes, Emily Chang, Philip J. Uren

Writing-Original Draft: Vladimir Gainullin, Frank Diehl, Tomasz M. Beer

Writing-Review and Editing: Vladimir Gainullin, Melissa Gray, Madhav Kumar, Stephen Luebker, Amy Lehman, Omair Choudhry, Julia Roberta, Darl Flake, Avinash Shanmugam, Kevin Cortes, Emily Chang, Philip J. Uren, Amin Mazloom, Jorge Garces, Tomasz M. Beer, Frank Diehl

Project Administration: Amin Mazloom, Jorge Garces, Frank Diehl

Conflicts of Interest

V.G.G., **M.G.**, **M.K., S.L., A.M.L., O.A.C., J.R., D.D.F. II, A.S., K.C., E.C., P.J.U., J.G., A.M., T.M.B.**, and **F.D.** are employees at Exact Sciences Corporation and declare stock ownership at Exact Sciences Corporation. **A.M., J.G.**, **T.M.B.**, and **F.D.** hold leadership roles at Exact Sciences Corporation. **P.J.U.** declares stock ownership at Biora Therapeutics outside this work.

T.M.B. declares stock ownership at Osteologic and Osheru, and a consulting/advisory role at AstraZeneca outside this work. **G.A.S.** declares advisory board service at Candel Therapeutics, consulting fees from Freenome and Biodesix, and research funding (Institutional) from NIH (RO1 screening in frail populations), Nucleix, Inc, Biodesix, Delfi Diagnostics, and Freenome outside this work. **G.A.S.** serves on the American Cancer Society roundtable steering committee outside this work. **D.W.C.** declares no conflicts of interest. **R.W.G.** declares speaker's bureau service at Bayer and Johnson & Johnson, research funding from Francis Medical, MDX Health, Levee, and Dendreon; and travel reimbursement from Francis Medical outside this work.

References

1 Siegel RL, Kratzer TB, Wagle NS et al. Cancer statistics, 2026. CA: A Cancer Journal for Clinicians 2026; 76 (1): e70043.

2 Curry SJ, Krist AH, Owens DK et al. Screening for Cervical Cancer: US Preventive Services Task Force Recommendation Statement. JAMA 2018; 320 (7): 674-686.

3 Davidson KW, Barry MJ, Mangione CM et al. Screening for Colorectal Cancer: US Preventive Services Task Force Recommendation Statement. JAMA 2021; 325 (19): 1965-1977.

4 Potter AL, Bajaj SS, Yang CJ. The 2021 USPSTF lung cancer screening guidelines: a new frontier. Lancet Respir Med 2021; 9 (7): 689-691.

5 Siu AL. Screening for Breast Cancer: U.S. Preventive Services Task Force Recommendation Statement. Ann Intern Med 2016; 164 (4): 279-296.

6 NORC (University of Chicago). Only 14% of Cancers Are Detected Through a Preventive Screening Test.; 2022. https://www.norc.org/content/dam/norc-org/pdfs/State-Specific%20PCDSs%20chart%201213.pdf

477 7 Imperiale TF, Porter K, Limburg PJ. Stool and Blood DNA Tests for Colorectal Cancer
478 Screening. Reply. N Engl J Med 2024; 390 (23): 2225.
479 8 Kim JJ, Burger EA, Regan C et al. Screening for Cervical Cancer in Primary Care: A
480 Decision Analysis for the US Preventive Services Task Force. JAMA 2018; 320 (7): 706-
481 714.
482 9 Lehman CD, Arao RF, Sprague BL et al. National Performance Benchmarks for Modern
483 Screening Digital Mammography: Update from the Breast Cancer Surveillance
484 Consortium. Radiology 2017; 283 (1): 49-58.
485 10 Pinsky PF, Gierada DS, Black W et al. Performance of Lung-RADS in the National Lung
486 Screening Trial: a retrospective assessment. Ann Intern Med 2015; 162 (7): 485-491.
487 11 Braunstein GD, Ofman JJ. Criteria for evaluating multi-cancer early detection tests.
488 2021.
489 12 White T, Algeri S. Estimating the lifetime risk of a false positive screening test result.
490 PLOS ONE 2023; 18 (2): e0281153.
491 13 Bao H, Yang S, Chen X et al. Early detection of multiple cancer types using
492 multidimensional cell-free DNA fragmentomics. Nat Med 2025.
493 14 Gainullin V, Bae J, Guthrie VB et al. Abstract A056: Performance of multi-biomarker
494 class reflex testing in a prospectively-collected cohort. Clinical Cancer Research 2024;
495 30 (21_Supplement): A056-A056.
496 15 Gainullin V, Manesse M, Zhuang F et al. Performance of a multi-target, multi-cancer
497 early detection (MCED) blood test in a prospectively collected cohort. medRxiv 2025:
498 2025.2008.2024.25334244.
499 16 Gao Q, Lin YP, Li BS et al. Unintrusive multi-cancer detection by circulating cell-free
500 DNA methylation sequencing (THUNDER): development and independent validation
501 studies. Ann Oncol 2023.

502 17 Lennon AM, Buchanan AH, Kinde I et al. Feasibility of blood testing combined with PET-
503 CT to screen for cancer and guide intervention. Science 2020; 369 (6499): eabb9601.
504 18 Schrag D, Beer TM, McDonnell CH, 3rd et al. Blood-based tests for multicancer early
505 detection (PATHFINDER): a prospective cohort study. Lancet 2023; 402 (10409): 1251-
506 1260.
507 19 Nabavizadeh N, McDonnell C, III, Kurbegov D et al. LBA64 Safety and performance of a
508 multi-cancer early detection (MCED) test in an intended-use population: Initial results
509 from the registrational PATHFINDER II study. Annals of Oncology 2025; 36: S1720.
510 20 Buchanan AH, Lennon AM, Choudhry OA et al. Multiyear Clinical Outcomes of Cancers
511 Diagnosed Following Detection by a Blood-Based Multicancer Early Detection Test.
512 Cancer Prev Res (Phila) 2024; 17 (8): 349-353.
513 21 Cohen JD, Li L, Wang Y et al. Detection and localization of surgically resectable cancers
514 with a multi-analyte blood test. Science 2018; 359 (6378): 926-930.
515 22 Klein EA, Richards D, Cohn A et al. Clinical validation of a targeted methylation-based
516 multi-cancer early detection test using an independent validation set. Ann Oncol 2021;
517 32 (9): 1167-1177.
518 23 SEER Explorer: An interactive website for SEER cancer statistics [Internet]. Surveillance
519 Research Program, National Cancer Institute (2023). [Accessed March 4, 2025].
520 https://seer.cancer.gov/statistics-network/explorer/
521 24 Kisiel JB, Ebbert JO, Taylor WR et al. Shifting the Cancer Screening Paradigm:
522 Developing a Multi-Biomarker Class Approach to Multi-Cancer Early Detection Testing.
523 Life 2024; 14 (8): 925.
524 25 Tyson C, Li KH, Cao X et al. Tumor localization strategies of multi-cancer early detection
525 tests: a quantitative assessment. JNCI Cancer Spectrum 2025.
526 26 American Cancer Society. Cancer Facts and Figures, 2023. Atlanta, GA: American
527 Cancer Society. [Accessed March 4, 2025]. https://www.cancer.org/content/dam/cancer-

528 org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2023/2023-
529 cancer-facts-and-figures.pdf
530 27 Chhatwal J, ElHabr A, Tyson C et al. Correlation of unobserved incidence of cancer in
531 earlier stages with the observed incidence. Journal of Clinical Oncology 2023; 41
532 (16_suppl): 10634-10634.
533 28 Katerov VE, Fleming HL, Rugowski DE et al. The detection of multiple cancer types with
534 an extended set of methylation and protein markers. Journal of Clinical Oncology 2023;
535 41 (16_suppl): 3040-3040.
536 29 U.S. Census Bureau 2020. Profile of General Population and Housing Characteristics.
537 Accessed May 2, 2025].
538 https://data.census.gov/table?g=010XX00US&d=DEC+Demographic+Profile

539

540
541
542
543
544
545
546
547
548
549
550
551
552
553
554

555 Table 1. Independent mini-holdout classifier selection test set performance comparison

556

	MP V1 Classifier Target specificity ≥98.0% (95% CI), n/N	**MP V1 Classifier Target specificity ≥97.0%** (95% CI), n/N		**MP V2 Classifier Target specificity ≥97.0%** (95% CI), n/N		**MP V2 Classifier Target specificity ≥97.0%** (95% CI), n/N	
Measured Specificity	99.4% (98.3, 99.8), 506/509	99.0% (97.7, 99.6), 504/509		99.0% (97.7, 99.6), 504/509		99.0% (97.7, 99.6), 504/509	
		Versus MP V1 classifier @ ≥98% Target Specificity		Versus MP V1 classifier @ ≥98% Target Specificity		Versus MP V1 classifier @ ≥97% Target Specificity	
		Absolute Increase*	Relative Increase**	Absolute Increase*	Relative Increase**	Absolute Increase*	Relative Increase**
Overall Sensitivity	60.9% (51.6, 69.5), 67/110	3.6%	5.9%	11.8%	19.4%	8.2%	12.7%
Sensitivity by Stage:							
Stage I	31.8% (16.4, 52.7), 7/22	4.6%	14.5%	9.1%	28.6%	4.5%	12.4%
Stage II	63.6% (43.0, 80.3), 14/22	4.6%	7.2%	9.1%	14.3%	4.5%	6.6%
Stage III	68.2% (47.3, 83.6), 15/22	9.1%	13.3%	18.2%	26.7%	9.1%	11.8%
Stage IV	95.7% (79.0, 99.2), 22/23	0.0%	0.0%	4.3%	4.5%	4.3%	4.5%
Unknown	42.9% (24.5, 63.5), 9/21	0.0%	0.0%	19.0%	44.3%	19.0%	44.3%
Stages I-II	47.7% (33.8, 62.1), 21/44	4.6%	9.6%	9.1%	19.1%	4.5%	8.6%

557 *Absolute Increase=Sensitivity B% - Sensitivity A%. ** Relative Increase=Sensitivity B% –
558 Sensitivity A%/Sensitivity A% x 100.

559

560

561

562

563

564

565

566

567

568

569

570

571

572 Table 2. MP V1 and V2 classifier performance in the test set

573

	MP V1 Classifier Performance (95% CI)	**MP V2 Classifier Performance (95% CI)**	**p-value ‡**	**MP V2 Classifier Versus MP V1 Classifier[a]**	
Specificity (n=2,434)	98.5% (97.9, 98.9)	97.4% (96.7, 97.9)	<0.0001 ****	Absolute Increase[b]	Relative Increase[c]
Overall Sensitivity, all cancers (n=729)	50.9% (47.3, 54.5)	57.8% (54.1, 61.3)	<0.0001 ****	6.9%	13.6%
Stage I (n=182)	15.4% (10.9, 21.3)	21.4% (16.1, 27.9)	0.003 **	6.0%	39.0%
Stage II (n=163)	38.0% (30.9, 45.7)	46.0% (38.5, 53.7)	0.00024 ***	8.0%	21.1%
Stage III (n=180)	67.8% (60.6, 74.2)	76.1% (69.4, 81.8)	<0.0001 ****	8.3%	12.2%
Stage IV (n=172)	85.5% (79.4, 90.0)	89.5% (84.1, 93.3)	0.039 *	4.0%	4.7%
Unknown Stage (n=32)	37.5% (22.9, 54.7)	50.0% (33.6, 66.4)	0.125 n.s.	12.5%	33.3%
Stages I/II (n=345)	26.1% (21.7, 31.0)	33.0% (28.3, 38.2)	<0.0001 ****	6.9%	26.4%
Overall Sensitivity, breast and prostate cancers excluded (n=590)	56.8% (52.8, 60.7)	64.1% (60.1, 67.8)	<0.0001 ****	7.3%	12.9%
Stage I (n=145)	17.2% (12.0, 24.2)	24.8% (18.5, 32.4)	0.003 **	7.6%	44.2%
Stage II (n=109)	48.6% (39.4, 57.9)	57.8% (48.4, 66.6)	0.002 **	9.2%	18.9%
Stage III (n=151)	73.5% (66.0, 79.9)	81.5% (74.5, 86.8)	0.00049 ***	8.0%	10.9%
Stage IV (n=155)	86.5% (80.2, 91.0)	90.3% (84.6, 94.0)	0.070 n.s.	3.8%	4.4%
Unknown Stage (n=30)	40.0% (24.6, 57.7)	53.3% (36.1, 69.8)	0.125 n.s.	13.3%	33.3%
Stages I/II (n=254)	30.7% (25.4, 36.6)	39.0% (33.2, 45.1)	<0.0001 ****	8.3%	27.0%
Overall Sensitivity[d]	54.8% (50.4, 59.2)	61.6% (57.2, 65.8)	<0.0001 ****	6.8%	12.4%
Overall Sensitivity[e]	63.7% (58.0, 69.0)	67.8% (62.2, 72.9)	0.002**	4.1%	6.4%

574 [a]MP V2 classifier (97.4% specificity) versus an MP V1 classifier (98.5% specificity). [b]Absolute
575 Increase=Sensitivity MP V2% - Sensitivity MP V1%. [c]Relative Increase=Sensitivity MP V2% –
576 Sensitivity MP V1%/Sensitivity MP V1% x 100. [d]Sensitivity excluding cancer organ types with
577 average-risk standard of care screening (i.e., excluding breast, prostate, cervix, colon and
578 rectum). [e]Sensitivity for the 6 most aggressive cancer organ types with the shortest 5-year
579 survival rate reported by National Cancer Institute Surveillance, Epidemiology, and End Results
580 (SEER) (i.e., pancreas, esophagus, liver, lung and bronchus, stomach, and ovary). ‡, McNemar
581 test; n.s., not significant; *, significant; **moderately significant; ***very significant; **** highly
582 significant.

583

584 Table 3. MP V2 clinical validation set participant demographics and clinical characteristics
585

	Characteristic	Cancer n (%)	Non-Cancer n (%)	Total N (%)
Sex	Female	160 (49.4%)	415 (51.9%)	575 (51.2%)
	Male	164 (50.6%)	385 (48.1%)	549 (48.8%)
Age (years)	Mean (SD)	66.9 (8.2)	63.7 (8.9)	64.6 (8.8)
	(Min, Max)	(50, 84)	(50, 84)	(50, 84)
Race	White	284 (87.7%)	657 (82.1%)	941 (83.7%)
	Black or African American	30 (9.3%)	91 (11.4%)	121 (10.8%)
	Asian	2 (0.6%)	10 (1.3%)	12 (1.1%)
	American Indian or Alaska Native	2 (0.6%)	6 (0.8%)	8 (0.7%)
	Native Hawaiian or Other Pacific Islander	1 (0.3%)	1 (0.1%)	2 (0.2%)
	More than one race	1 (0.3%)	6 (0.8%)	7 (0.6%)
	Unknown	4 (1.2%)	29 (3.6%)	33 (2.9%)
Ethnicity	Hispanic/Latino	11 (3.4%)	155 (19.4%)	166 (14.8%)
	Not Hispanic/Latino	302 (93.2%)	636 (79.5%)	938 (83.6%)
	Unknown	11 (3.4%)	7 (0.9%)	18 (1.6%)
	Missing	0	2	2
BMI (kg/m^2)	Mean (SD)	29.4 (7.1)	30.2 (6.7)	30.0 (6.8)
	(Min, Max)	(14.2, 61)	(16.4, 66.4)	(14.2, 66.4)
	Missing	1	0	1
Cigarette Smoking History	Current	50 (15.5%)	118 (14.8%)	168 (15.0%)
	Former	141 (43.8%)	215 (26.9%)	356 (31.7%)
	Never	131 (40.7%)	467 (58.4%)	598 (53.3%)
	Missing	2	0	2

Alcohol Consumption History	Current	153 (47.5%)	411 (51.6%)	564 (50.4%)
	Former	55 (17.1%)	95 (11.9%)	150 (13.4%)
	Never	114 (35.4%)	291 (36.5%)	405 (36.2%)
	Missing	2	3	5
Family History of Cancer in 1st Degree Relative	No	105 (32.4%)	363 (45.4%)	468 (41.7%)
	Yes	208 (64.2%)	424 (53.1%)	632 (56.3%)
	Unknown	11 (3.4%)	12 (1.5%)	23 (2.0%)
	Missing	0	1	1

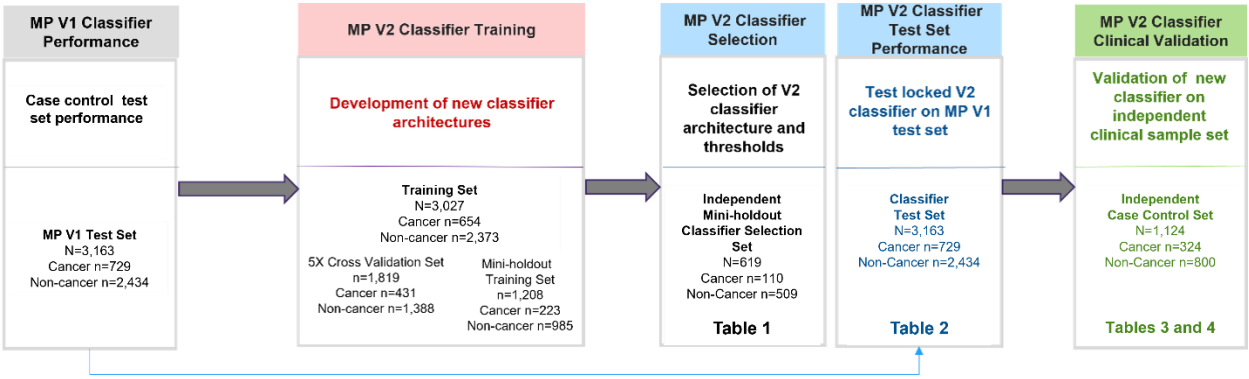
BMI, body mass index kg/m^2

606 Table 4. MP V2 classifier validation set performance

(n/N)	Performance	95% CI
Specificity (779/800)	97.4%	(96.0, 98.3)
Overall Sensitivity, all cancers (134/324)	41.4%	(36.1, 46.8)
Stage I (17/106)	16.0%	(10.3, 24.2)
Stage II (26/83)	31.3%	(22.4, 41.9)
Stage III (37/70)	52.9%	(41.3, 64.1)
Stage IV (54/65)	83.1%	(72.2, 90.3)
Stages I-II (43/189)	22.8%	(17.4, 29.2)
Overall Sensitivity, breast and prostate cancers excluded (124/233)	55.6%	(49.0%, 62.0)
Stage I (15/56)	26.8%	(17.0, 39.6)
Stage II (24/56)	42.9%	(30.8, 55.9)
Stage III (35/55)	63.6%	(50.4, 75.1)
Stage IV (50/56)	89.3%	(78.5, 95.0)
Stages I-II (39/112)	34.8%	(26.6%, 44.0)
Overall Sensitivity* (66/126)	52.4%	(43.7, 60.9)
Overall Sensitivity** (72/106)	67.9%	(58.5, 76.0)

607 *Sensitivity excluding cancer organ types with average-risk standard of care screening (i.e.,
608 excluding breast, prostate, cervix, colon and rectum). **Sensitivity for the 6 most aggressive
609 cancer organ types with the shortest 5-year survival rate reported by National Cancer Institute
610 Surveillance, Epidemiology, and End Results (SEER) (i.e., pancreas, esophagus, liver, lung and
611 bronchus, stomach, and ovary).
612
613
614
615

616 Figure 1. MP V2 classifier development workflow



617

618

619

620 MP, Methylation-Protein.

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639