ACS NANO

www.acsnano.org

ARTICLE

Noninvasive Precision Screening of Prostate Cancer by Urinary Multimarker Sensor and Artificial Intelligence Analysis

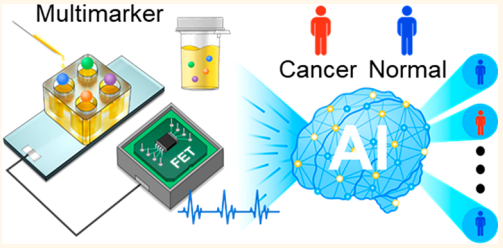
Hojun Kim,[†] Sungwook Park,[†] In Gab Jeong,[†] Sang Hoon Song, Youngdo Jeong, Choung-Soo Kim, and Kwan Hyi Lee*

Cite This: https://dx.doi.org/10.1021/acsnano.0c06946

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Screening for prostate cancer relies on the serum prostate-specific antigen test, which provides a high rate of false positives (80%). This results in a large number of unnecessary biopsies and subsequent overtreatment. Considering the frequency of the test, there is a critical unmet need of precision screening for prostate cancer. Here, we introduced a urinary multimarker biosensor with a capacity to learn to achieve this goal. The correlation of clinical state with the sensing signals from urinary multimarkers was analyzed by two common machine learning algorithms. As the number of biomarkers was increased, both algorithms provided a monotonic increase in screening performance. Under the best combination of biomarkers, the machine learning algorithms screened prostate cancer patients with more than 99% accuracy using 76 urine specimens. Urinary multimarker biosensor leveraged by machine learning analysis can be an important strategy of precision screening for cancers using a drop of bodily fluid.



KEYWORDS: *cancer screening, urine, multimarker, artificial intelligence, prostate cancer*

About 1.3 million men worldwide were diagnosed with prostate cancer (PCa) in 2018.[1] PCa is the second most common cancer in males, excluding non-melanoma skin cancer.[2] The current standard PCa screening methods include the monitoring of serum prostate-specific antigen (PSA) levels and the digital rectal examination (DRE), in which doctors check for abnormalities, such as lumps or hard areas, on the prostate tissue.[3] The DRE causes considerable discomfort and is less accurate than the PSA test; however, it increases biomarker concentration and some researchers recognize its helpful role in assessing PCa in males who have a PSA level in the "gray range" (2.5−4.0 ng/mL).[3] In contrast, the PSA test is a stand-alone screening method. In the United States, more than 25 million men receive a PSA test each year,[4] and approximately 1 million men have a biopsy as a result of an elevated PSA level.[5] Ultimately, less than 20% of those who undergo a biopsy are diagnosed with PCa.[6,7] These statistics show clearly that the PSA test has a very high rate of false positives (80%). These false results have led to a large number of unnecessary and invasive biopsies and subsequent overtreatment, which causes considerable side effects, including erectile dysfunction and urinary incontinence.[8] Therefore, it is critical that a noninvasive,[9−11] DRE-free PCa screening system is developed[12] that is highly accurate and minimizes false results.

Considering the large number of the invasive serum PSA test, strategies for increased patient compliance would bring a broad impact to society. In this aspect, naturally voided urine is a particularly useful specimen due to high patient compliance. Moreover, there are urinary protein databases available,[13] suggesting that urine is a useful specimen for cancer screening. However, the low biomarker concentration imposes a challenge in utilizing urine as a PCa screening test.[14] Therefore, a sensor that detects trace amounts of biomarkers in urine has great potential for PCa screening.

There has been a surge in research to find an alternative biomarker that can surpass the screening performance of the PSA test in PCa screening.[15] However, PCa screening systems based solely on a single reported biomarker can result in low diagnostic accuracy because simultaneously seeking a specific and sensitive biomarker is challenging.[16,17] In other words, each single biomarker has a limited correlation with the clinical

Received: August 18, 2020
Accepted: December 2, 2020

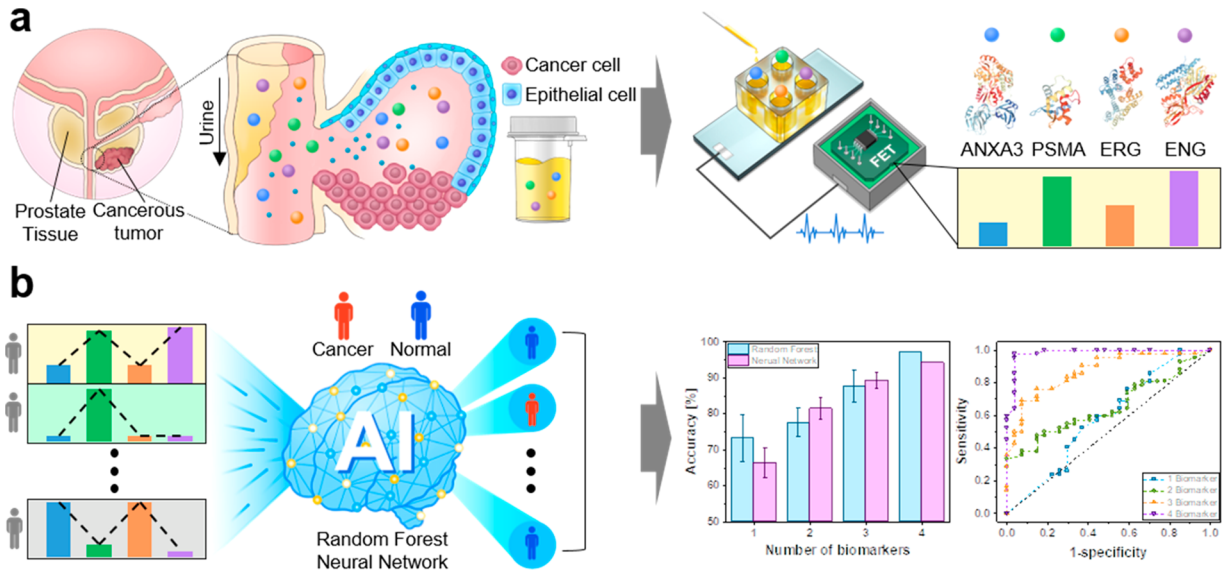


Figure 1. Overview of a ML-assisted multimarker biosensing system for PCa screening using a drop of urine and its clinical performance. (a) Biomarkers diffused passively from prostate tissue to the urethra and were collected from naturally voided urine. Urine was collected in a sterilized specimen cup containing 10 vol % RNA stabilizer and 1 vol % antibiotics and stored in a refrigerator for less than a week before shipping the sample to the laboratory. To measure the electrical signals from the four different biomarkers in the urine, the four sensing channels in the EG of an FET biosensor were conjugated to antibodies, thus capturing different biomarkers. Without any pretreatment, urine was added directly to the four different sensing channels. After 20 min of reaction time in the four channels, the bottom gate voltage shifts were measured at the reference current (1 nA). Therefore, each urine sample generated four independent sensing signals from the corresponding biomarkers, as shown in the schematics. (b) The set of sensing signals collected for each patient were then analyzed using ML to screen the patient for PCa. Seventy-six urine samples were measured three times, thereby generating 912 biomarker signals or 228 sets of sensing signals. We used RF and NN algorithms to analyze the multimarker signals. Both algorithms provided an increased accuracy, and the AUROC increased in size as the number of biomarkers was increased.

output of PCa. In this regard, a combination of multiple biomarkers is a promising approach[18,19] to overcoming the limitations of relying on a single biomarker in PCa screening.[20] To get benefits from the multimarker approach[21−23] in biosensor-based PCa screening, it is first necessary to design biomarker panels that can improve diagnostic accuracy.[24,25] For example, a highly sensitive biomarker can be coupled with a highly specific biomarker to construct a highly accurate PCa screening system. In a multimarker-based biosensor system, a set of sensing data from each biomarker panel can be obtained, which can then translate into a clinical decision for an individual. However, the multimarker system has multiple independent variables, such as the biomarker-sensing signal profiles and their intensities. As a result, as the number of markers in the system increases, the process of converting the sensing data into diagnostic output becomes exponentially complicated. Therefore, a statistical approach is necessary.

In this study, we developed a urinary multimarker sensor, able to measure trace amounts of biomarkers from naturally voided urine. The relation of a clinical outcome with a set of sensing signals from four different PCa biomarkers was analyzed by two different machine learning (ML) algorithms. Throughout the study, our multimarker sensing platform demonstrated that precision screening for PCa using a drop of urine is feasible.

RESULTS AND DISCUSSION

Urine was selected as the specimen for screening patients who may require a biopsy. Urine is of particular interest for PCa screening because the urethra is surrounded by prostate tissue, which allows for the passive diffusion of biomarkers from the PCa cells into the urine.[26] Obtaining a urine specimen is also noninvasive, and hence it leads to increased patient compliance. However, the low concentration of biomarkers creates a challenge when using urine for translational research.[14,27] We used a highly sensitive dual-gate field-effect transistor (DGFET) as a urinary multimarker sensor to deal with this challenge. This DGFET is composed of a disposable four-channel extended gate (EG), which is separated from its transducer to improve sensing performance and reliability. In our previous studies,[28−30] the DGFET biosensor enabled the detection of a trace number of protein-based biomarkers from patients and animal specimens in 20 min. To extract clinically significant information from complicated multimarker-sensing signals, we introduced two ML algorithms (random forest (RF) and neural network (NN)). Specifically, RF and NN were compared to find the best algorithm and combination of biomarkers that provided the highest accuracy in PCa screening.

The general research outline is shown in Figure 1. Multiple biomarkers from PCa cells diffuse passively into the urethra and are transported during urination. The naturally voided urine was collected from potential patients with PCa. The specimens were applied to the DGFET biosensor to obtain electrical signals from the different biomarkers. In this study, four biomarker signals were obtained: annexin A3 (ANXA3), prostate-specific membrane antigen (PSMA), erythroblast transformation-specific related gene protein (ERG), and endoglin (ENG). Obtained data from 76 urine specimens were partitioned randomly into a training data set (70% of total) and a test data set (30% of total). After the training phase, ML screened the patients with PCa in the test set. The PCa screening performance improved monotonically as more

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX

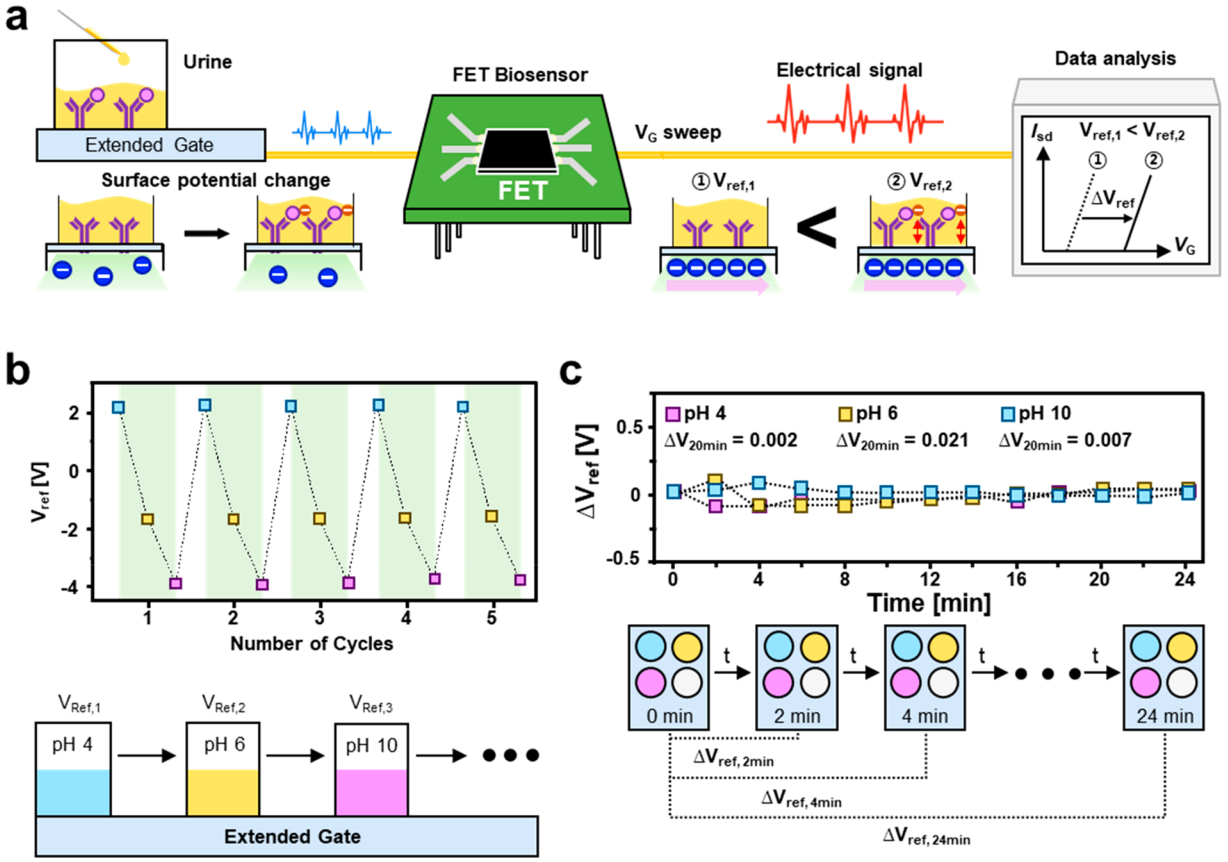


Figure 2. The working principle of urine measurements and the pH stability of the biosensor sensing performance. (a) Principle of urine measurement. Urine was first added to the sensing channels and then the biomarkers in the urine bound with the surface-conjugated antibodies, which resulted in a potential change to the surface. Before and after the binding event, the bottom gate voltage was swept, and the source−drain current was recorded. As the biomarker binding events modulated the surface potential, the threshold voltage shifted. The bottom gate voltage at the reference current (1 nA) was specifically used to calculate the voltage shift. The voltage shifts for all urine biomarkers were recorded and used for further analyses. (b) Three buffer solutions, at pH 4, 6 (normal pooled urine), and 10, were added to the same sensing channel sequentially, and the corresponding bottom gate voltage at the reference current (1 nA) was measured. After five cycles of the solution exchange process, small voltage shifts (0.13, 0.095, and 0.029 V for pH 4, 6, and 10, respectively) were observed, suggesting a robust sensing performance in the urine environment. (c) The shift in the bottom gate voltage was measured over 24 min in the same buffer solutions as in b. The voltage shifts were small (0.002, 0.021, and 0.007 V for pH 4, 6, and 10, respectively), further confirming the device's stability in the urine environment.

biomarkers were combined. The results suggest that PCa screening using a drop of urine is almost 100% accurate.

Choice of Biomarkers. Our primary focus in relation to biomarker selection was on their clinical relation with PCa. We specifically selected the following four pathophysiologically uncorrelated biomarkers that are upregulated in patients with PCa: ANXA3, PSMA, ERG, and ENG. ANXA3 is found in urine with high specificity to PCa,[31] and ANXA3 expression has been observed in prostatic intraepithelial neoplasia, a precursor to PCa.[32] PSMA is a membrane protein expressed on the apical surface of endothelial cells. PSMA overexpression is associated with a higher PCa grade and androgen deprivation, suggesting that PSMA plays a functional role in PCa progression.[33] In the case of ERG, its expression is associated strongly with an increased level of TMPRSS2/ERG fusion genes, by which tumor progression and invasion are promoted.[34] Lastly, increased ENG expression is found in prostate endothelial cells during tumor angiogenesis and inflammation. ENG is also expressed in prostatic intraepithelial neoplasia, suggesting its potential as a prognostic marker.[35] Considering the working principle of a field-effect transistor (FET), it is necessary to test whether or not chosen biomarkers can induce enough charge polarization and subsequent threshold voltage change in the DGFET system. The detailed sensing characteristics of the four biomarker candidates were evaluated, and these are represented in Figure S1. All proposed biomarkers show a linear relationship between voltage shift and concentration level. The details of the sensing characteristics of these biomarkers in the literature are shown in Supplementary Table 1.

In this study, untreated urine was added directly to each sensing channel of the EG. The principle of urine measurement is shown in Figure 2a. The binding of an antigen with an antibody resulted in a change in the surface potential. This change was monitored by the shift of the bottom gate voltage at the reference current (1 nA). The voltage shifts for all biomarkers were recorded and used for further analysis. In our previous work,[30] we successfully measured blood-based samples containing a significantly higher amount of protein than urine. Unlike the constant pH of blood (pH 7.35−7.45),[36] however, the pH of urine varies significantly from person to person (pH 4.5−8.0),[37] which may affect the sensing characteristics. To eliminate this issue, the pH stability of the DGFET was tested using five consecutive cycles of pH 4, 6

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX

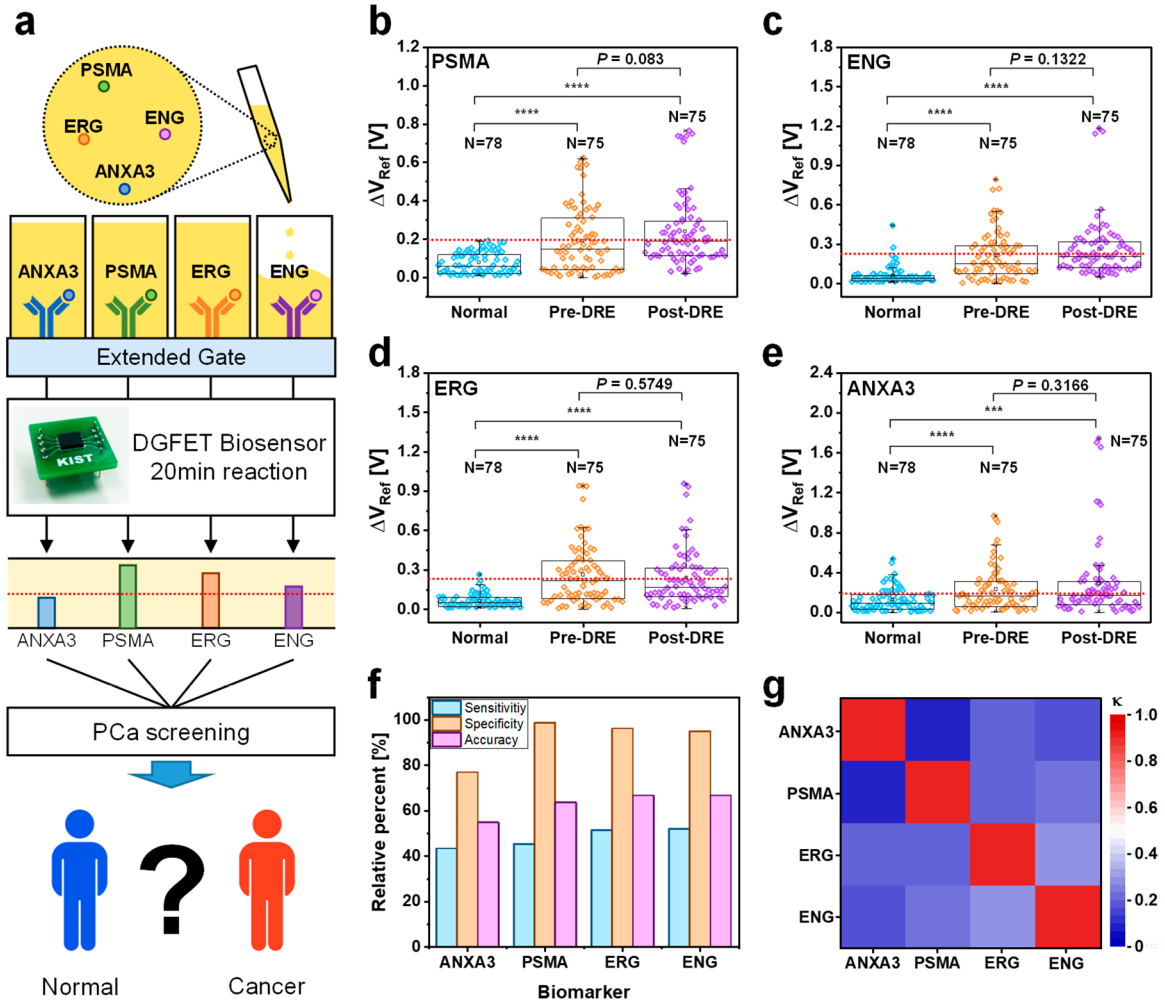


Figure 3. Raw sensing data from 76 naturally voided urine specimens and single-marker analyses. (a) Experimental workflow of 76 urine measurements. (b–e) Each urine specimen was measured from three independent panels with four biomarker channels. A cutoff signal was chosen from the voltage shift induced by normal pooled urine (0.19 V). Box plots are shown for (b) PSMA, (c) ENG, (d) ERG, and (e) ANXA3 biomarkers. Statistical significance was determined by unpaired Student's *t* tests (*$p < 0.001$ ****$p < 0.0001$). (f) Sensing performance of the single biomarkers. For the PSMA, ENG, and ERG biomarkers, the sensitivity, specificity, and accuracy were similar. ANXA3 had a much lower specificity (77%) than the other biomarkers due to the high number of false positives, which suggests that it had the lowest sensing performance of the four. (g) Correlation matrix of biomarker pairs. From the graph, the ERG + ENG pairing was found to have the highest correlation ($\kappa = 0.27$), while the ANXA3 + PSMA pairing had the lowest ($\kappa = 0.04$).**

(normal pooled urine), and 10. Figure 2b shows the bottom gate voltages at the reference current during the five cycles of pH change in the same channel of EG. After five cycles, the voltage shift showed small hysteresis (changes of 0.13, 0.095, and 0.029 V for pH 4, 6, and 10, respectively), proving that the sensor performance was stable in the required sample reaction environment. Figure 2c shows the small voltage shift for the different pH solutions (0.002, 0.021, and 0.007 V for pH 4, 6, and 10, respectively) over 24 min, further confirming the stability of the device performance.

Patient Choice. We divided urine samples into three groups: normal, PCa with pre-DRE, and PCa with post-DRE. Benign prostatic hyperplasia (BPH) is similar to PCa in that it causes the prostate gland to become enlarged. Considering that BPH is a common disease in men over the age of 40, having BPH in the normal group is ideal; however, it is challenging because the levels of certain biomarkers are affected.[38] In this study, we included urine specimens from patients with BPH in the normal group so that we could screen for PCa accurately. The naturally voided urine specimens were collected from the normal group during a regular health checkup. In the case of patients with PCa, urine was collected twice, before and after the DRE procedure. The sample population of the Gleason score is shown in Figure S2. Supplementary Table 2 shows a summary of the information for the 51 patients studied. The differences in age and serum PSA levels between the normal and PCa groups were statistically significant. On the other hand, the difference in prostate volume was not significant. For unbiased prediction, none of these factors are accounted for in the ML algorithm.

Single-Biomarker Analysis. The raw sensing data and single-biomarker analysis from 76 naturally voided urine samples are shown in Figure 3. Figure 3a shows the experimental scheme. Each urine sample was added to the four sensing channels, which were conjugated with four different antibodies. After 20 min, the biosensor measured the gate voltage shift at the reference current (1 nA). Therefore, each specimen provided four sensing data points, an analysis of which resulted in accurate PCa screening. The sensing data from all urine specimens are shown as box plots in Figure 3b–

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX

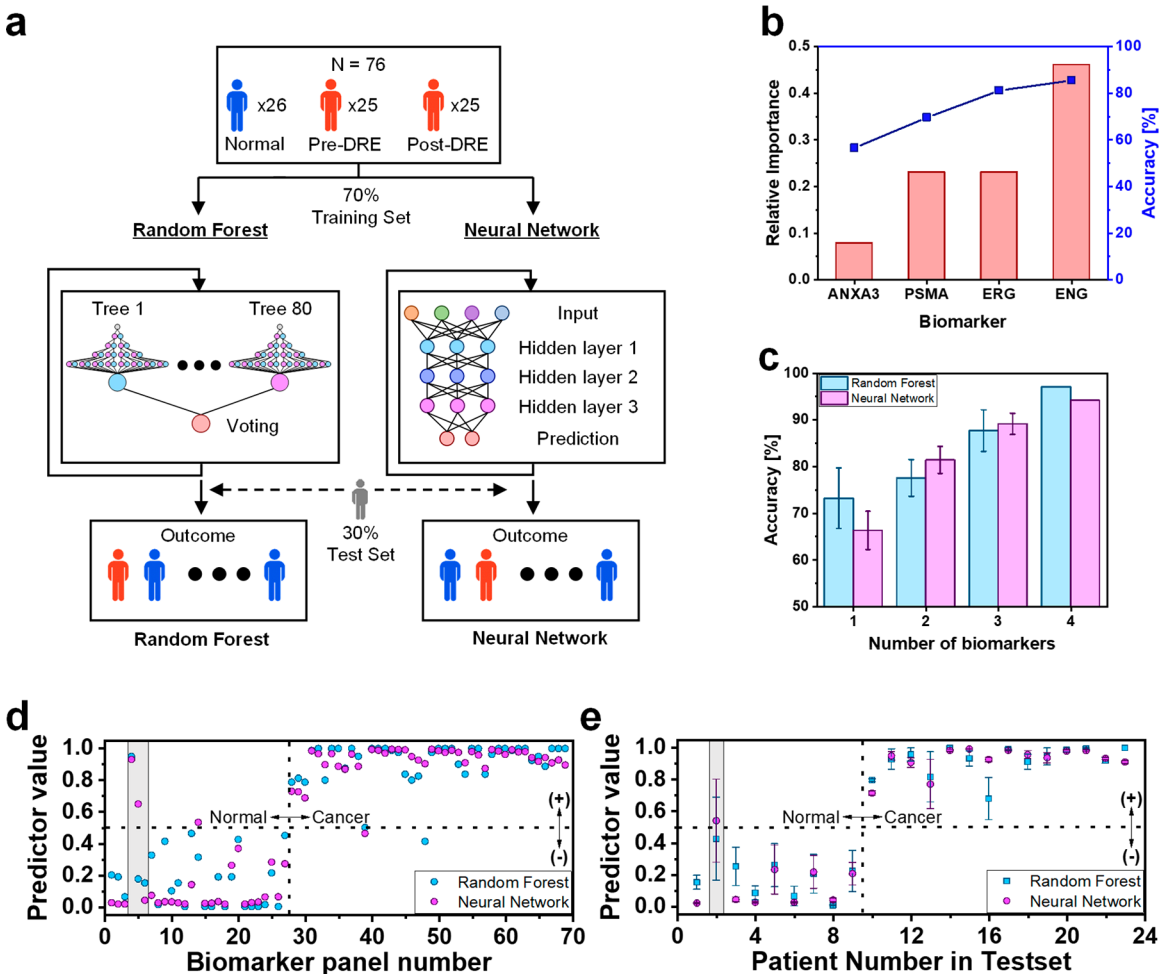


Figure 4. ML analysis of the multimarker system. (a) RF and NN algorithms used in this study. A total of 76 clinical samples were separated randomly into a training set (70%) and a test set (30%). After the learning process with a training set, the performance of each algorithm was evaluated by the test set using 23 specimens. (b) Sensing characteristics of each biomarker provided by the RF method. Interestingly, although the PSMA, ERG, and ENG biomarkers had a similar sensing performance, they showed a notably different importance in their decision processes. ANXA3 showed the least importance (7.9%), which was consistent with our previous analysis. (c) Average accuracy predicted by both algorithms with different numbers of biomarkers. NN had a higher accuracy when two or three biomarkers were combined. For both algorithms, an increased number of biomarkers in the panel increased the accuracy monotonically, which confirms that the multimarker approach assisted by ML provides superior screening results than the single-biomarker system, which missed more than half of the patients with PCa at high specificity. (d) The predictor value of each biomarker panel was provided by both algorithms. For the normal group (panel numbers 1−27), RF and NN had one and three biomarker panels, respectively, which exceeded a threshold of 0.5. In the cancer groups, both algorithms predicted one panel with a predictor value of less than 0.5. It should be noted that three consecutive panels used the same urine. The gray box represents the region where more than two consecutive panels showed a false signal. (e) Predictor value based on individuals. Patient number 2 showed a false signal from NN, resulting in 95.7% accuracy. RF provided 100% accuracy at the patient level. Error bars represent s.d.

e (averaged biological replications are shown in Figure S3). Each specimen was measured three times to count any sensing signal variations. A cutoff voltage value was chosen from the shift induced by the normal pooled urine (0.19 V). In Figure 3b, signals from the PSMA biomarker are shown. Signals from pre- and post-DRE urine specimens are similar and distinct from the normal group, which barely passed the cutoff value. In the case of the normal group, only one (0.19 V) of the 78 data points reached the cutoff voltage shift, while the cancer (pre- and post-DRE) groups had 30 and 38 of the 75 total data points exceed the threshold, respectively. The *p*-values between normal and either pre- or post-DRE were less than 0.0001, suggesting that the signals are distinguishable. However, the *p*-value from the pre- and post-DRE groups was calculated as 0.083, indicating that the two data sets are not statistically significant. Therefore, the PSMA sensing results suggest that our direct urine sensing can screen patients with PCa at a similar level of accuracy, regardless of the DRE procedure. Similar results were obtained from the ENG and ERG biomarkers (Figures 3c−d and S3). In contrast, the ANXA3 biomarker showed a different result (Figure 3e). ANXA3 had the lowest sensing performance among the four biomarkers. A significantly large proportion of the data points in the normal group (18 out 78) had a larger voltage shift than the cutoff value. In contrast, the numbers of data points that exceeded 0.19 V in the pre- and post-DRE groups were similar to those for the other biomarkers (34 and 35 for the pre- and post-DRE groups, respectively). Importantly, the interquartile spacing of the box plots overlapped considerably between the normal and cancer groups, indicating that the ANXA3 biomarker shows

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX

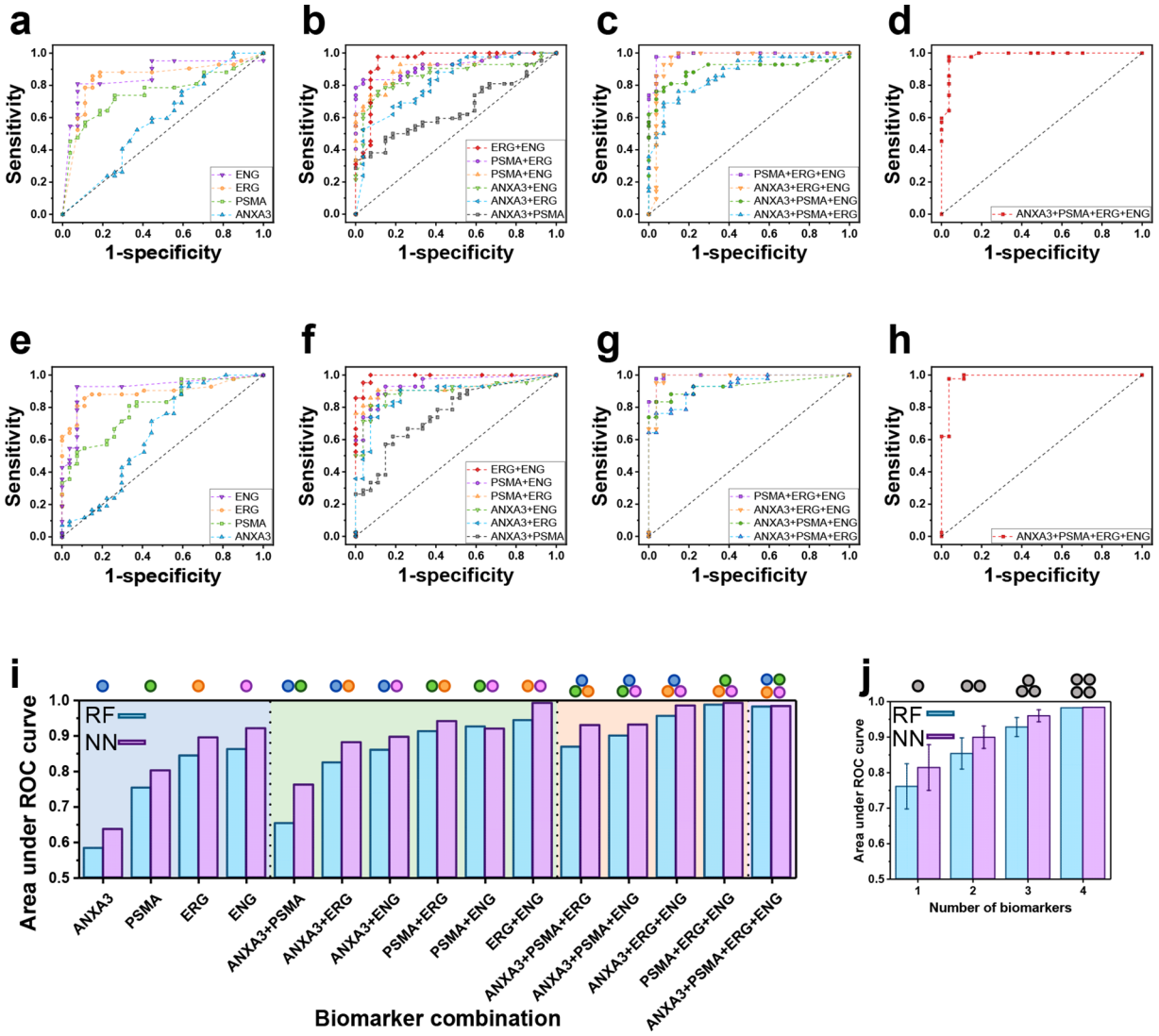


Figure 5. Sensing performance of ML algorithms for PCa screening with different biomarker combinations. (a–h) ROC curves of the multimarker sensing systems with different biomarker combinations at the biomarker panel level. The ROC curves from the RF analysis, with 1 (a), 2 (b), 3 (c), and 4 (d) biomarkers, and those from the NN analysis, with 1 (e), 2 (f), 3 (g), and 4 (h) biomarkers, indicate that a better sensing performance (points closer to each axis) can be achieved with an increased number of combined biomarkers. (i) The AUROC, which represents the sensing performance, is the single most important parameter. For single-biomarker systems, the areas under the ROC curves for ANXA3, PSMA, ERG, and ENG were in the range of 0.58–0.86 and 0.64–0.92 for RF and NN, respectively. When these biomarkers were combined, the areas under the ROC curves could reach 0.99 for both algorithms with the best combination of biomarkers. This suggests that it is critical to ascertain the optimal biomarker combination. (j) Average AUROC with different numbers of biomarkers. Again, the increasing trend was observed, confirming the importance of a multimarker approach. Error bars represent s.d.

the least sensing performance as a single biomarker in our DGFET biosensor.

The sensitivity, specificity, and accuracy of the four biomarkers used in this study are shown in Figure 3f. The sensitivity (the ability to correctly identify patients with the disease) was similar for all biomarkers, in the range of 43.3–52.0%, where the minimum sensitivity was achieved by the ANXA3 biomarker. PSMA, ERG, and ENG had more than 94.9% specificity, whereas ANXA3 had only 76.9%. The low sensing performance of ANXA3 was consistent with the box plot in Figure 3, where the sensing signals from all three groups largely overlapped. Lastly, three of the biomarkers, PSMA, ERG, and ENG, achieved a similar accuracy (63.6–66.7%), while ANXA3 achieved only 54.8% accuracy. Overall, the test outcomes indicate that the screening performance of the PSMA, ERG, and ENG biomarkers was similar, while the performance of the ANXA3 biomarker lagged behind the others. Considering the main purpose of the study, reducing the rate of false positive results is of primary interest because that would lead to a reduction in the number of unnecessary biopsies. In this regard, all four biomarkers already outperform the PSA test. However, higher specificity always compromises sensitivity. In our test, we missed approximately half of the patients with PCa at this high level of specificity. In other words, the current single biomarker tests can minimize unnecessary biopsies; however, they can miss a significant portion of patients with PCa because of a high rate of false negatives. Therefore, it is necessary to formulate a strategy to increase specificity without compromising sensitivity.

To achieve this challenging goal with a multimarker approach, it is necessary to choose uncorrelated biomarkers. A lower correlation between biomarkers will increase the

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX–XXX

information content.[39] Figure 3g shows the Pearson correlation coefficient matrix of biomarker pairs. ERG and ENG were found to be the most correlated (κ = 0.27), while the ANXA3−PSMA pair showed the least correlation (κ = 0.04). In fact, the average correlation factors of the individual biomarkers were 0.13, 0.15, 0.22, and 0.22 for ANXA3, PSMA, ERG, and ENG, respectively. The low κ values for all pairs of biomarkers suggest that they are not correlated and have distinct sensing outcomes.[40] In other words, the lists of individuals requiring biopsies, which are decided by each biomarker, do not overlap significantly. Therefore, the data with a low κ value suggest that there is an opportunity to enhance the sensing performance if the sensing signals from each biomarker are blended properly.

Multimarker Analysis by ML. We applied two common ML algorithms (RF[41] and NN[42]) to illustrate the importance of multimarker approach (Figure 4a). Both algorithms were trained by 70% of the total data set under supervised learning and validated its performance by screening PCa from a blinded test set (30% of the total data set). The reference voltage shifts from each biomarker were input into both algorithms, giving two different types of output: normal and PCa. It should be noted here that the cutoff value (0.19 V) was not considered in the algorithm. This means that both algorithms learn the raw data of the DGFET to consider even minor characteristics caused by urine that otherwise would be ignored.

RF is useful in terms of understanding variables. For example, RF can provide useful information, such as the importance of each biomarker in its decision process. In Figure 4b the relative importance of each biomarker in the four-biomarker combination system is provided by RF. Interestingly, the PSMA, ERG, and ENG biomarkers that showed a similar sensing performance had a notably different importance in the RF algorithm. PSMA and ERG showed similar importance (0.23 for both biomarkers), but the accuracy obtained from RF was higher in ERG, indicating that the importance does not necessarily follow each biomarker's accuracy. As expected, from the calculations in Figure 3f ANXA3 showed the least importance (only 7%), which suggests that some biomarkers barely contribute to diagnostic performance.[43]

Figure 4c shows the average accuracy of PCa screening by RF and NN using different numbers of biomarkers at the biomarker panel level. For both algorithms, on average, the accuracy increased monotonically as the number of biomarkers in the panel increased. On average, the accuracy of the RF and NN algorithms with a single biomarker was 73.2% and 66.3%, respectively. As the number of biomarkers increased, the average accuracy improved dramatically, to 97.1% and 94.2% for RF and NN, respectively. The improving average accuracy following an increase in the number of biomarkers in two different ML algorithms shows clearly that the multimarker approach provides a far more accurate clinical output, whereas the single-biomarker analysis missed about half of the PCa patients at this high specificity. This result supports the idea that pathologically independent biomarker combination broadens the disease-specific information.

In general, ML decisions are based on the predictor value. The closer the predictor value is to either 0 or 1, the greater the certainty of the algorithm's prediction. The predictor values of each biomarker panel at four-biomarker combinations are shown in Figure 4d. In terms of biomarker panels, the RF predictions were all correct, except for one in each of the normal and cancer biomarker panels. The NN algorithm missed three biomarker panels from the normal group and one from the cancer group. As each urine sample was measured three times, the three consecutive panel numbers corresponded to each specimen. For NN, one specimen had two consecutive incorrect panels (panel numbers 4 and 5), which is highlighted in gray. The average predictor value of the individual is shown in Figure 4e. The accuracy of the RF algorithm was 100%, meaning that there were zero false results in the PCa screening. With the NN algorithm, one false positive specimen was located in the gray box. It should be noted that even though RF prediction was perfect in this data set, there were samples with predictor values close to 0.5, suggesting that the prediction certainty was less than 100%. Therefore, it is important to set an uncertainty window in the diagram, with patients who fall within this window taking additional tests. To validate our analysis in this limited number of data sets, four more predictions with different validation sets were performed and showed less than 5% accuracy variations (Figure S10).

The receiver operating characteristic (ROC) curve represents the performance of our ML models.[44] Figure 5a−d shows the ROC curves of the biomarker panels predicted by the RF algorithms with different numbers and combinations of biomarkers. As the number of biomarkers in the system increases, the ROC curves get closer to each axis, meaning that there is an increase in accuracy. In addition, the increased number of biomarkers in the system leads to decreased variations in the ROC curves. The same trend was observed with the NN algorithm (Figure 5e−h).

The single parameter representing the model's sensing performance is area under ROC curve (AUROC). The AUROCs for the systems in Figure 5a−h are shown in Figure 5i, which confirms that an increased number of biomarkers leads to a greater AUROC. For single-biomarker systems, the AUROC is the largest with ENG, followed by ERG, PSMA, and ANXA3. When these biomarkers were combined, AUROC varied significantly depending on the biomarker combination. This variation is significant, as it is often greater than the effect of the number of biomarkers. For example, ENG + ERG has a higher sensing performance than PSMA + ERG + ANXA3. RF and NN are complicated algorithms with nonlinear functions. This means that increasing the number of markers does not necessarily guarantee a monotonic increase in screening accuracy.[43] In fact, the inclusion of ANXA3 in our multimarker system often resulted in a lower sensing performance for both the RF and NN algorithms. This indicates that there are optimal biomarker combinations. For example, the best sensing output in terms of AUROC was achieved from ENG + ERG + PSMA and ENG + ERG for RF and NN, respectively. In general, the AUROC is slightly larger for NN than for RF, but the trend was the same. The average AUROC *versus* the number of biomarkers is shown in Figure 5j. As demonstrated in this figure, the areas under the ROC curves increase monotonically as the number of biomarkers is increased. NN shows a better model performance than RF in terms of the ROC curves on average, but this is not statistically significant. The same trend was observed from the ROC curves based on individuals (Figure S5).

Clinical Analysis of False Signals. Studying patients with false signals can provide an insight into the relationship between the ML prediction and the disease status of a patient, which can be used to improve the ML algorithm. The

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX

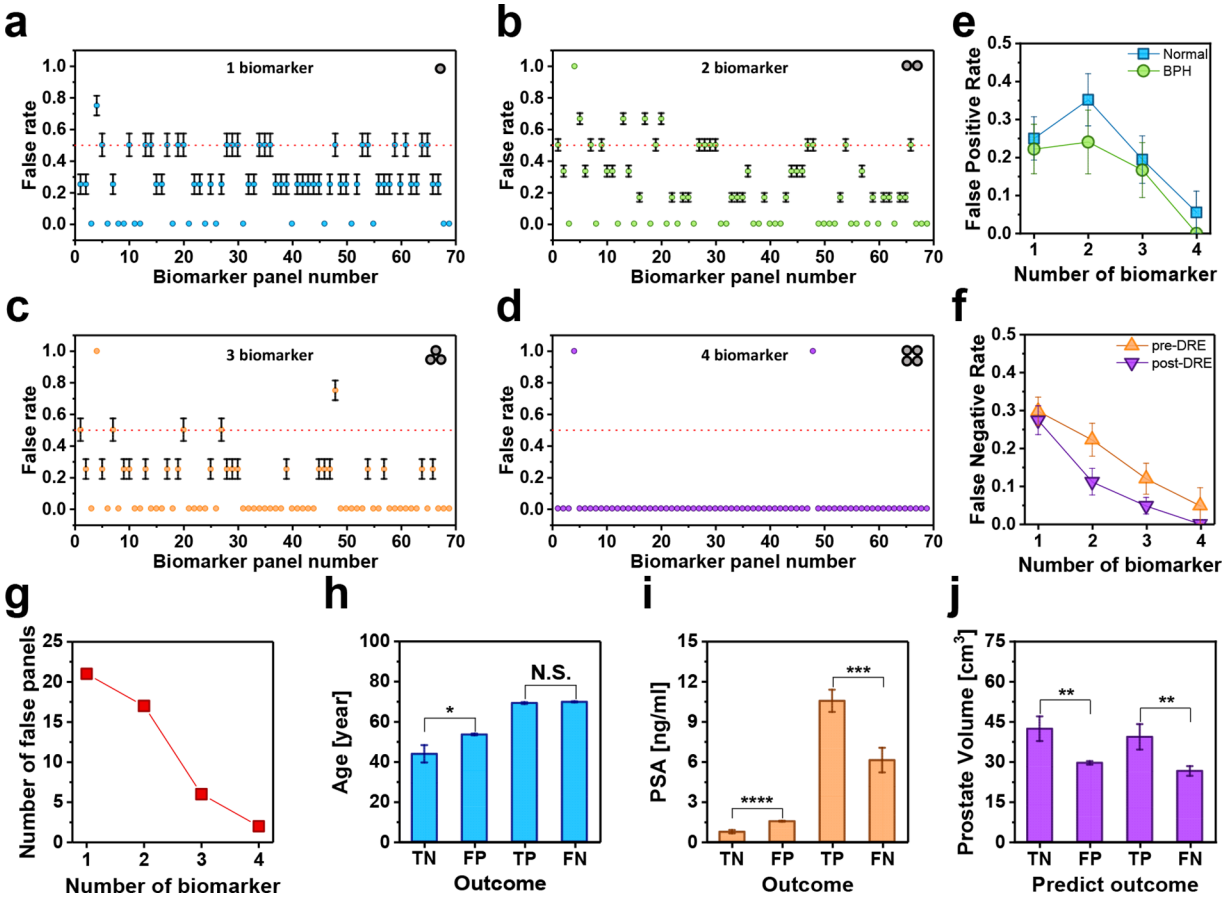


Figure 6. Clinical analysis of false signals predicted by RF. (a, b, d, e) Scatter plots of false rates with different numbers of biomarker combinations in the test set (total of 69 panels in test set). When four biomarkers were combined, only two panels gave a false signal (d). (e) False positive rate of biomarker panels for normal individuals and patients with BPH. RF had a better screening performance in patients with BPH than in normal individuals without BPH. (f) False negative rate of biomarker panels for pre- and post-DRE groups. The RF predictions in the post-DRE group were more accurate than those in the pre-DRE group for all biomarker combinations. However, the pre-DRE group had 95.2% sensitivity with the combination of four biomarkers. (g) Total number of biomarker panels with a false rate higher than 0.5. As the number of biomarkers in the panel was increased, the number of false panels decreased from 21 to 2. (h–j) Clinical analysis of outcomes from all biomarker combinations. (h) The age difference between true negative and false positive is statistically significant, indicating that elderly people have a higher chance of testing positive for PCa than younger people in the normal group. (i) PSA had a much higher correlation in both the normal and cancer groups. In each group, individuals testing positive have a higher PSA than those testing negative. (j) The prostate volume showed the same trend for both normal and cancer groups. Patients with a small prostate volume had a higher chance of receiving a false screening. Statistical significance was determined by unpaired Student's *t* tests (*$p < 0.001$, ****$p < 0.0001$). Error bars represent s.d.**

information can also be used to construct a feedback loop for the patients, whereby individuals showing high false rates can take additional tests to ensure that they do not have PCa. Figure 6a, b, c, and d show the RF-predicted average false rate of each biomarker panel in the test set with different numbers of biomarkers. The scatter plots show clearly that an increased number of biomarkers reduces the overall false rate. The four-biomarker combination results in Figure 6d identify only two false panels. Interestingly, these two panels always have a nonzero false rate, regardless of the number of biomarkers. Figure 6g shows the number of panels with a false rate higher than 0.5 at different numbers of biomarkers. As the number of biomarkers is increased, the number of false panels decreases (21, 17, 6, and 2 panels for one-, two-, three-, and four-biomarker combinations).

Figure 6e shows the false positive rates of the normal and BPH groups with different numbers of biomarkers, which indicates that our sensing data do not highlight any statistically meaningful differences between BPH and normal individuals, except when two biomarkers are combined ($p = 0.03$). The false negative rates of the pre- and post-DRE cancer groups are represented (Figure 6f); these rates were higher in the pre-DRE group than in the post-DRE group, indicating that the post-DRE urine specimens in our study had better specificity. However, both groups had more than 95.2% sensitivity in four-biomarker combinations.

To study the factors affecting the false signals in our ML algorithms, clinical analyses of the confusion matrix are shown in Figure 6h–j. It is noteworthy that no clinical parameters were considered in ML algorithms, eliminating concerns regarding bias induced by clinical parameters. In the normal group, elderly people were found to have a higher chance of being incorrectly screened for PCa than younger individuals; however, there was no statistical significance in the cancer groups. In contrast, the serum PSA level differences were significant; those with higher serum PSA levels had a higher chance of testing positive in both normal and cancer groups. Lastly, people with a lower prostate volume had a higher

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX–XXX

chance of receiving a false screening result for both the normal and cancer groups.

CONCLUSIONS

In this study, we developed urinary multimarker sensors and AI analysis for precision screening for PCa, using a drop of naturally voided urine. Specifically, four pathophysiologically uncorrelated biomarkers, PSMA, ENG, ERG, and ANXA3, were used. A total of 76 naturally voided urine specimens from healthy and PCa-diagnosed individuals were measured directly using a DGFET biosensor, comprising four biomarker channels conjugated to antibodies capturing each biomarker. A single-biomarker analysis provided the average accuracy of only 62.9%, and it missed approximately half of the patients with PCa.

In our ML-assisted multimarker sensing approach, the two different ML algorithms (RF and NN) were applied. The average accuracy of both algorithms was monotonically increased as more biomarkers were included in the ML algorithms. However, having more biomarkers did not necessarily guarantee an improved sensing performance. For example, the inclusion of ANXA3 actually decreased the accuracy in most cases. In addition, constructing the biomarker panel using the minimum number of biomarkers is necessary owing to the limited resources available; therefore, it is important to choose synergistic biomarkers. One of the key factors to consider when designing a biomarker panel is the importance of each biomarker provided by RF. In this regard, RF is a better algorithm than NN because it provides parameters that allow an understanding of the variables.[45] At the best biomarker combinations, RF showed 100% accuracy in 23 individuals, or 97.1% accuracy in terms of panels, in a blinded test set regardless of the DRE procedure.

When individuals showing false signals were analyzed, a strong correlation between RF prediction and the clinical data was observed. As only biomarker signals were considered in the RF algorithm, the correlation provides additional insight on understanding predictor value. Importantly, RF prediction follows the prevalent view of the community studying prostate cancer. In the normal group, elderly people have a higher chance of being diagnosed with PCa. In the case of serum PSA level, individuals with a higher serum PSA level are prone to testing positive for PCa in both normal and cancer groups. Lastly, those with a small prostate volume had a higher chance of receiving a false screening in each group, presumably due to the low biomarker concentration, whereby the corresponding signal confuses the RF decision process. This correlation further supports the reliability of the sensing platform.

Our PCa screening platform showed exceptional prostate cancer screening performance using 76 urine specimens. However, our study is limited to randomly participating Korean males. This leaves the opportunity for additional experiments with different races. Another useful study to validate our approach is measuring a large number of samples with controlled age, PSA level, and prostate volume. An effort to look for more efficient biomarkers and their combinations is also necessary.

With increased patient numbers, the accumulation of data will drive an evolution of our ML algorithm in the multimarker sensing platform. Therefore, after validation with a large amount of data, our multimarker sensing platform has the potential to shift the current PCa screening paradigm. On-site and accurate PCa screening in 40 min with a drop of naturally voided urine will directly impact the human healthcare system in the foreseeable future.

EXPERIMENTAL METHODS

Materials. Human PSMA protein (orb245517) and anti-PSMA antibodies (orb25921) were purchased from Biorbyt. ENG protein (H00002022-P01), anti-ENG antibodies (H00002022-M01), ANXA3 protein (H00000306-P01), and anti-ANXA3 antibodies (H00000306-M12) were purchased from Abnova. ERG protein (ab185597) and anti-ERG antibodies (ab133264) were purchased from Abcam. Glutaraldehyde and 3-aminopropyltriethoxysilane (APTES) were purchased from Sigma-Aldrich. Bovine serum albumin (BSA, 97061-416) was obtained from VWR International. Normal urine (991-03-P-1) was acquired from Lee Biosolutions. All chemicals and proteins were used without further purification steps.

Urine Specimen Collection. The research protocol was approved by the Institutional Review Board of the Asan Medical Center (Seoul, Republic of Korea), and informed consent was obtained from all subjects (protocol no. 2014-0959). All clinical results and urine samples were obtained from patients after informed consent had been given, in accordance with the Asan Medical Center guidelines. The protocols for the human urine assay were approved by the Ethics Committee of the Asan Medical Center, and the human urine assays were performed in accordance with the approved guidelines.

Fabrication of a DGFET Sensor. The DGFET used in this study consisted of a FET and an EG. The manufacture of the DGFET device was described in our previous paper.[30] In brief, a six-inch p-type (100) silicon wafer with a 750 nm buried oxide layer and a 100 nm top silicon layer was used. Inductively coupled plasma reactive-ion etching (ICP-RIE) and photolithography were applied to define the active region. A-10 nm-thick oxide layer on the top gate was deposited using a dry oxidation process. On this layer, a Ti/TiN/Al/TiN multilayer was grown *via* a sputtering system, following ICP-RIE, to form the gate electrode. The arsenic ion implantation formed the drain and source parts.

An EG with four polydimethylsiloxane (PDMS) channels was fabricated on a glass substrate[46] with an indium tin oxide (ITO, 300 nm) film on top. The four PDMS channels, with a surface area of 0.5 cm^2, were prepared using a hole punch (0.8 cm in diameter). These four channels were attached to the ITO substrate through plasma bonding after O_2 plasma treatment (70 W, 30 sccm of O_2 gas flow for 1 min, Plasma System Cute, Femto Science). Once attached, an annealing process was performed at 60 °C for 10 min.

Surface Modification of the Extended Gate. At the surface of the EG, the hydroxyl groups were introduced by O_2 plasma treatment at 70 W for 1 min. Then, amine groups were introduced by dipping the surface in fresh 5 vol % APTES (ethanol) for 1 h at room temperature.[47] The APTES-treated surface was then washed with ethanol three times followed by sonication with ethanol for 1 min to remove unreacted APTES. Afterward, the APTES layers were cured at 120 °C for 30 min using a hot plate. To conjugate the APTES layer with antibodies or proteins through an amide bond, it was treated with a 2.5 wt % glutaraldehyde solution (20 mM HEPES buffer (pH 7.5), 1 h at room temperature). At the same time, the biomarker conjugation solutions of 5 μg/mL are prepared in 1× phosphate-buffered saline (PBS; pH 7.4). Subsequently, 100 μL of the four biomarkers (PSMA, ANXA3, ENG, and ERG) solutions were added into each channel and reacted for 1 h at room temperature. After washing with 1× PBS (pH 7.4) three times, 0.2 M ethanolamine was added for 1 h to inactivate the unreacted aldehyde groups. Then, the EG was incubated with 3 wt % BSA solution in 1× PBS (pH 7.4) for 1 h at room temperature to block nonspecific binding. After the BSA incubation, EG was washed with fresh 1× PBS (pH 7.4) three times.

Measurement Setup of the DGFET Biosensor. A dual-channel parameter analyzer (4200A-SCS, Keithley) was used to obtain an I_D-V_G curve on the DGFET. For the I_D-V_G curve, the sweeping voltage of the bottom gate was −10 to 10 V and the drain voltage was 3 V. To obtain the standard curve, the voltage shift at various concentrations

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX

of PSMA, ANXA3, ENG, and ERG (0.1 fg/mL to 10 ng/mL) in normal urine was measured. The initial voltage signal was obtained from the I_D-V_{BG} curves in a 1× PBS solution. The solution was then removed, and the urine specimens were injected directly into the modified EG and allowed to react for 20 min. After washing with a 1× PBS solution three times, the washed solution was removed from the channels. Finally, a clean 1× PBS solution was added and the final voltage signal was measured. The detection voltage (voltage shift) was obtained by subtracting the final (t = 20 min) and the initial (t = 0 min) voltage signals at the reference current (1 nA).

For multimarker sensing, the same procedure was used, except that the four channels in the EG were conjugated with different antibodies. The initial voltage signal was recorded in a 1× PBS solution, followed by simultaneous urine incubation at each channel for 20 min, and then washed three times using a 1× PBS solution. Finally, a clean 1× PBS solution was added and the final voltage signal was measured.

Statistical Analysis for Multimarker Biosensors. To generate predictor values for each biomarker panel, custom-built ML software was developed by Python 3.6.1, with multiple open libraries, including Scikit Learn, Panda, Numpy, and Matplotlib. RF decision trees and feedforward neural networks for NN algorithms were used in this study. RF comprises 80 random decision trees with binary classifications for cancer and normal cases, with a majority-voted outcome being collected for all biomarker panels. For NN, a feedforward neural network with three hidden layers of three nodes was used. The NN model was implemented using Keras with a TensorFlow framework. To prevent an overfitting issue, we used the early stop regularization technique by optimizing hyperparameters. For both algorithms, a supervised learning method was used, and they were iteratively trained by randomly assigning 70% of the total data set. The rest of the blinded test set (30% of total) was then used to validate the screening performance of the algorithms. The voltage shifts from the four biomarker sensing channels were input into both RF and NN algorithms, to evaluate the signal combinations and generate the predictor value. The predictor value of 0.5 was used as a threshold (≥0.5 for cancer; <0.5 for normal) for the biomarker panels in calculating the accuracy. As each urine sample was measured three times, an average of the predictor values was used to make a final decision on the PCa screening (averaged standard deviation is less than 0.042 and 0.066 for RF and NN, respectively). As more biosensing data are accumulated, the performance of the algorithm will become even more enhanced, thus further validating its effectiveness.

Statistical Analysis. Each clinical specimen was tested three independent times, and the data were expressed as the mean ± standard deviation, unless otherwise indicated. Statistical comparisons between two groups were based on unpaired Student's t tests. Differences were considered significant at $p < 0.05$.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsnano.0c06946.

Reference voltage shift at varied biomarker concentrations, sample population based on Gleason score, averaged sensing data from 76 urine specimens, our DGFET sensor have a potential to screen prostate cancer patient without DRE, prediction characteristics of ML algorithms at the patient level, false rates predicted by the RF in the test set at the patient level, false rates predicted by the NN in the test set at the biomarker panel level, false rates predicted by the NN in the test set at the patient level, average false rates at different number of biomarker combinations, diagnostic performances of four more validation sets, I_D-V_G curves at each stage of surface functionalization, selected urinary biomarkers for PCa from the literature, clinical characteristics and patient demographics of the PCa screening (PDF)

AUTHOR INFORMATION

Corresponding Author

Kwan Hyi Lee − *Biomaterials Research Center, Korea Institute of Science and Technology (KIST), Seoul 02792, Republic of Korea; KU-KIST Graduate School of Converging Science and Technology, Korea University, Seoul 02841, Republic of Korea;* orcid.org/0000-0002-7891-844X; Email: kwanhyi@kist.re.kr

Authors

Hojun Kim − *Biomaterials Research Center, Korea Institute of Science and Technology (KIST), Seoul 02792, Republic of Korea;* orcid.org/0000-0001-7974-1044

Sungwook Park − *Biomaterials Research Center, Korea Institute of Science and Technology (KIST), Seoul 02792, Republic of Korea*

In Gab Jeong − *Department of Urology, Asan Medical Center (AMC), University of Ulsan College of Medicine, Seoul 05505, Republic of Korea*

Sang Hoon Song − *Department of Urology, Asan Medical Center (AMC), University of Ulsan College of Medicine, Seoul 05505, Republic of Korea*

Youngdo Jeong − *Biomaterials Research Center, Korea Institute of Science and Technology (KIST), Seoul 02792, Republic of Korea*

Choung-Soo Kim − *Department of Urology, Asan Medical Center (AMC), University of Ulsan College of Medicine, Seoul 05505, Republic of Korea*

Complete contact information is available at: https://pubs.acs.org/10.1021/acsnano.0c06946

Author Contributions

†H. Kim, S. Park, and I. G. Jeong contributed equally.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was supported by the Basic Science Research Program through the National Research Foundation (NRF) funded by the Ministry of Science and ICT (MSIT) (2020R1A2C2009291) of Republic of Korea. This work was also supported by the Korea Medical Device Development Fund grant funded by the Korean government (the Ministry of Science and ICT, the Ministry of Trade, Industry and Energy, the Ministry of Health & Welfare, Republic of Korea, the Ministry of Food and Drug Safety) (Project Number: 202012D07).

REFERENCES

(1) Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R. L.; Torre, L. A.; 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.

(2) Culp, M. B.; Soerjomataram, I.; Efstathiou, J. A.; Bray, F.; Jemal, A. Recent Global Patterns in Prostate Cancer Incidence and Mortality Rates. *Eur. Urol.* **2020**, *77*, 38−52.

(3) Wolf, A. M. D.; Wender, R. C.; Etzioni, R. B.; Thompson, I. M.; D'Amico, A. V.; Volk, R. J.; Brooks, D. D.; Dash, C.; Guessous, I.; Andrews, K.; DeSantis, C.; Smith, R. A. American Cancer Society

Guideline for the Early Detection of Prostate Cancer: Update 2010. *Ca-Cancer J. Clin.* **2010**, *60*, 70−98.
(4) Constantinou, J.; Feneley, M. R. PSA Testing: An Evolving Relationship with Prostate Cancer Screening. *Prostate Cancer Prostatic Dis.* **2006**, *9*, 6−13.
(5) Welch, H. G.; Fisher, E. S.; Gottlieb, D. J.; Barry, M. J. Detection of Prostate Cancer *via* Biopsy in the Medicare−SEER Population during the PSA Era. *JNCI J. Natl. Cancer Inst.* **2007**, *99*, 1395−1400.
(6) Siegel, R. L.; Miller, K. D.; Jemal, A. Cancer Statistics, 2020. *Ca-Cancer J. Clin.* **2020**, *70*, 7−30.
(7) Guo, D. P.; Thomas, I.-C.; Mittakanti, H. H.; Shelton, J. B.; Makarov, D. V.; Skolarus, T. A.; Cooperberg, M. R.; Sonn, G. A.; Chung, B. I.; Brooks, J. D.; Leppert, J. T. The Research Implications of Prostate Specific Antigen Registry Errors: Data from the Veterans Health Administration. *J. Urol.* **2018**, *200*, 541−548.
(8) Gori, D.; Banerjee, I.; Chung, B.; Ferrari, M.; Rucci, P.; Blayney, D.; Brooks, J. Extracting Patient-Centered Outcomes from Clinical Notes in Electronic Health Records: Assessment of Urinary Incontinence After Radical Prostatectomy. *eGEMs* **2019**, *7*, 43.
(9) Lin, Y.; Bariya, M.; Nyein, H. Y. Y.; Kivimäki, L.; Uusitalo, S.; Jansson, E.; Ji, W.; Yuan, Z.; Happonen, T.; Liedert, C.; Hiltunen, J.; Fan, Z.; Javey, A. Porous Enzymatic Membrane for Nanotextured Glucose Sweat Sensors with High Stability toward Reliable Non-invasive Health Monitoring. *Adv. Funct. Mater.* **2019**, *29*, 1902521.
(10) Peng, G.; Tisch, U.; Adams, O.; Hakim, M.; Shehada, N.; Broza, Y. Y.; Billan, S.; Abdah-Bortnyak, R.; Kuten, A.; Haick, H. Diagnosing Lung Cancer in Exhaled Breath Using Gold Nanoparticles. *Nat. Nanotechnol.* **2009**, *4*, 669−673.
(11) Chen, X.; Gole, J.; Gore, A.; He, Q.; Lu, M.; Min, J.; Yuan, Z.; Yang, X.; Jiang, Y.; Zhang, T.; Suo, C.; Li, X.; Cheng, L.; Zhang, Z.; Niu, H.; Li, Z.; Xie, Z.; Shi, H.; Zhang, X.; Fan, M.; et al. Non-Invasive Early Detection of Cancer Four Years before Conventional Diagnosis Using a Blood Test. *Nat. Commun.* **2020**, *11*, 3475.
(12) Kang, B. J.; Jeun, M.; Jang, G. H.; Song, S. H.; Jeong, I. G.; Kim, C.-S.; Searson, P. C.; Lee, K. H. Diagnosis of Prostate Cancer *via* Nanotechnological Approach. *Int. J. Nanomed.* **2015**, *10*, 6555−6569.
(13) Consortium, T. U. UniProt: A Worldwide Hub of Protein Knowledge. *Nucleic Acids Res.* **2019**, *47*, D506−D515.
(14) Fredolini, C.; Meani, F.; Alex Reeder, K.; Rucker, S.; Patanarut, A.; Botterell, P. J.; Bishop, B.; Longo, C.; Espina, V.; Petricoin, E. F.; Liotta, L. A.; Luchini, A. Concentration and Preservation of Very Low Abundance Biomarkers in Urine, Such as Human Growth Hormone (HGH), by Cibacron Blue F3G-A Loaded Hydrogel Particles. *Nano Res.* **2008**, *1*, 502−518.
(15) Prensner, J. R.; Rubin, M. A.; Wei, J. T.; Chinnaiyan, A. M. Beyond PSA: The Next Generation of Prostate Cancer Biomarkers. *Sci. Transl. Med.* **2012**, *4*, 127rv3.
(16) Kim, Y.; Jeon, J.; Mejia, S.; Yao, C. Q.; Ignatchenko, V.; Nyalwidhe, J. O.; Gramolini, A. O.; Lance, R. S.; Troyer, D. A.; Drake, R. R.; Boutros, P. C.; Semmes, O. J.; Kislinger, T. Targeted Proteomics Identifies Liquid-Biopsy Signatures for Extracapsular Prostate Cancer. *Nat. Commun.* **2016**, *7*, 11906.
(17) Cheng, Z.; Choi, N.; Wang, R.; Lee, S.; Moon, K. C.; Yoon, S.-Y.; Chen, L.; Choo, J. Simultaneous Detection of Dual Prostate Specific Antigens Using Surface-Enhanced Raman Scattering-Based Immunoassay for Accurate Diagnosis of Prostate Cancer. *ACS Nano* **2017**, *11*, 4926−4933.
(18) Brikun, I.; Nusskern, D.; Freije, D. An Expanded Biomarker Panel for the Detection of Prostate Cancer from Urine DNA. *Exp. Hematol. Oncol.* **2019**, *8*, 13.
(19) Wang, Y.-Z.; Ji, S.-Y.; Xu, H.-Y.; Zhao, W.; Xu, J.-J.; Chen, H.-Y. Bidirectional Electrochemiluminescence Color Switch: An Application in Detecting Multimarkers of Prostate Cancer. *Anal. Chem.* **2018**, *90*, 3570−3575.
(20) A, M. D.; Eugene, B. Future of Biomarkers in Acute Coronary Syndromes. *Circulation* **2003**, *108*, 250−252.
(21) Im, H.; Shao, H.; Park, Y. Il; Peterson, V. M.; Castro, C. M.; Weissleder, R.; Lee, H. Label-Free Detection and Molecular Profiling of Exosomes with a Nano-Plasmonic Sensor. *Nat. Biotechnol.* **2014**, *32*, 490−495.
(22) Ko, J.; Bhagwat, N.; Yee, S. S.; Ortiz, N.; Sahmoud, A.; Black, T.; Aiello, N. M.; McKenzie, L.; O'Hara, M.; Redlinger, C.; Romeo, J.; Carpenter, E. L.; Stanger, B. Z.; Issadore, D. Combining Machine Learning and Nanofluidic Technology to Diagnose Pancreatic Cancer Using Exosomes. *ACS Nano* **2017**, *11*, 11182−11193.
(23) Joung, H.-A.; Ballard, Z. S.; Wu, J.; Tseng, D. K.; Teshome, H.; Zhang, L.; Horn, E. J.; Arnaboldi, P. M.; Dattwyler, R. J.; Garner, O. B.; Di Carlo, D.; Ozcan, A. Point-of-Care Serodiagnostic Test for Early-Stage Lyme Disease Using a Multiplexed Paper-Based Immunoassay and Machine Learning. *ACS Nano* **2020**, *14*, 229−240.
(24) Broza, Y. Y.; Zhou, X.; Yuan, M.; Qu, D.; Zheng, Y.; Vishinkin, R.; Khatib, M.; Wu, W.; Haick, H. Disease Detection with Molecular Biomarkers: From Chemistry of Body Fluids to Nature-Inspired Chemical Sensors. *Chem. Rev.* **2019**, *119*, 11761−11817.
(25) Akkilic, N.; Geschwindner, S.; Höök, F. Single-Molecule Biosensors: Recent Advances and Applications. *Biosens. Bioelectron.* **2020**, *151*, 111944.
(26) Eskra, J. N.; Rabizadeh, D.; Pavlovich, C. P.; Catalona, W. J.; Luo, J. Approaches to Urinary Detection of Prostate Cancer. *Prostate Cancer Prostatic Dis.* **2019**, *22*, 362−381.
(27) Afkarian, M.; Bhasin, M.; Dillon, S. T.; Guerrero, M. C.; Nelson, R. G.; Knowler, W. C.; Thadhani, R.; Libermann, T. A. Optimizing a Proteomics Platform for Urine Biomarker Discovery. *Mol. Cell. Proteomics* **2010**, *9*, 2195−2204.
(28) Jeun, M.; Park, S.; Kim, Y.; Choi, J.; Song, S. H.; Jeong, I. G.; Kim, C.-S.; Lee, K. H. Self-Normalized Detection of ANXA3 from Untreated Urine of Prostate Cancer Patients without Digital Rectal Examination. *Adv. Healthcare Mater.* **2017**, *6*, 1700449.
(29) Choi, J.; Jeun, M.; Yuk, S.-S.; Park, S.; Choi, J.; Lee, D.; Shin, H.; Kim, H.; Cho, I.-J.; Kim, S. K.; Lee, S.; Song, C. S.; Lee, K. H. Fully Packaged Portable Thin Film Biosensor for the Direct Detection of Highly Pathogenic Viruses from On-Site Samples. *ACS Nano* **2019**, *13*, 812−820.
(30) Jeun, M.; Lee, H. J.; Park, S.; Do, E.; Choi, J.; Sung, Y.-N.; Hong, S.-M.; Kim, S.-Y.; Kim, D.-H.; Kang, J. Y.; Son, H.-N.; Joo, J.; Song, E. M.; Hwang, S. W.; Park, S. H.; Yang, D.-H.; Ye, B. D.; Byeon, J.-S.; Choe, J.; Yang, S.-K.; et al. A Novel Blood-Based Colorectal Cancer Diagnostic Technology Using Electrical Detection of Colon Cancer Secreted Protein-2. *Adv. Sci.* **2019**, *6*, 1802115.
(31) Schostak, M.; Schwall, G. P.; Poznanović, S.; Groebe, K.; Müller, M.; Messinger, D.; Miller, K.; Krause, H.; Pelzer, A.; Horninger, W.; Klocker, H.; Hennenlotter, J.; Feyerabend, S.; Stenzl, A.; Schrattenholz, A. Annexin A3 in Urine: A Highly Specific Noninvasive Marker for Prostate Cancer Early Detection. *J. Urol.* **2009**, *181*, 343−353.
(32) Köllermann, J.; Schlomm, T.; Bang, H.; Schwall, G. P.; von Eichel-Streiber, C.; Simon, R.; Schostak, M.; Huland, H.; Berg, W.; Sauter, G.; Klocker, H.; Schrattenholz, A. Expression and Prognostic Relevance of Annexin A3 in Prostate Cancer. *Eur. Urol.* **2008**, *54*, 1314−1323.
(33) Bravaccini, S.; Puccetti, M.; Bocchini, M.; Ravaioli, S.; Celli, M.; Scarpi, E.; De Giorgi, U.; Tumedei, M. M.; Raulli, G.; Cardinale, L.; Paganelli, G. PSMA Expression: A Potential Ally for the Pathologist in Prostate Cancer Diagnosis. *Sci. Rep.* **2018**, *8*, 4254.
(34) Ayala, G.; Frolov, A.; Chatterjee, D.; He, D.; Hilsenbeck, S.; Ittmann, M. Expression of ERG Protein in Prostate Cancer: Variability and Biological Correlates. *Endocr.-Relat. Cancer* **2015**, *22*, 277−287.
(35) Fujita, K.; Ewing, C. M.; Chan, D. Y. S.; Mangold, L. A.; Partin, A. W.; Isaacs, W. B.; Pavlovich, C. P. Endoglin (CD105) as a Urinary and Serum Marker of Prostate Cancer. *Int. J. Cancer* **2009**, *124*, 664−669.
(36) Goldberg, M.; Green, S. B.; Moss, M. L.; Marbach, C. B.; Garfinkel, D. Computer-Based Instruction and Diagnosis of Acid-Base Disorders: A Systematic Approach. *JAMA* **1973**, *223*, 269−275.

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX

(37) Cook, J. D.; Strauss, K. A.; Caplan, Y. H.; Lodico, C. P.; Bush, D. M. Urine PH: The Effects of Time and Temperature after Collection. *J. Anal. Toxicol.* **2007**, *31*, 486−496.
(38) Lorenzo, G.; Hughes, T. J. R.; Dominguez-Frojan, P.; Reali, A.; Gomez, H. Computer Simulations Suggest That Prostate Enlargement Due to Benign Prostatic Hyperplasia Mechanically Impedes Prostate Cancer Growth. *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 1152−1161.
(39) McRae, M. P.; Simmons, G.; Wong, J.; McDevitt, J. T. Programmable Bio-Nanochip Platform: A Point-of-Care Biosensor System with the Capacity To Learn. *Acc. Chem. Res.* **2016**, *49*, 1359−1368.
(40) Cohen, J. A Coefficient of Agreement for Nominal Scales. *Educ. Psychol. Meas.* **1960**, *20*, 37−46.
(41) Choi, D.-H.; Kitchen, G.; Kim, J. S.; Li, Y.; Kim, K.; Jeong, I. c.; Nguyen, J.; Stewart, K. J.; Zeger, S. L.; Searson, P. C. Two Distinct Types of Sweat Profile in Healthy Subjects While Exercising at Constant Power Output Measured by a Wearable Sweat Sensor. *Sci. Rep.* **2019**, *9*, 17877.
(42) Shin, H.; Oh, S.; Hong, S.; Kang, M.; Kang, D.; Ji, Y.; Choi, B. H.; Kang, K.-W.; Jeong, H.; Park, Y.; Hong, S.; Kim, H. K.; Choi, Y. Early-Stage Lung Cancer Diagnosis by Deep Learning-Based Spectroscopic Analysis of Circulating Exosomes. *ACS Nano* **2020**, *14*, 5435−5444.
(43) Kim, S.; Jung, S.; Park, Y.; Lee, J.; Park, J. Effective Liver Cancer Diagnosis Method Based on Machine Learning Algorithm. *In 2014 7th International Conference on Biomedical Engineering and Informatics* **2014**, 714−718.
(44) Webb, G. I.; Ting, K. M. On the Application of ROC Analysis to Predict Classification Performance Under Varying Class Distributions. *Mach. Learn.* **2005**, *58*, 25−32.
(45) Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5−32.
(46) Jang, L.; Lee, J.; Razu, M. E.; Jensen, E. C.; Kim, J. Fabrication of PDMS Nanocomposite Materials and Nanostructures for Biomedical Nanosystems. *IEEE Trans. Nanobioscience* **2015**, *14*, 841−849.
(47) Cho, S.; Park, W.; Kim, D.-H. Silica-Coated Metal Chelating-Melanin Nanoparticles as a Dual-Modal Contrast Enhancement Imaging and Therapeutic Agent. *ACS Appl. Mater. Interfaces* **2017**, *9*, 101−111.

https://dx.doi.org/10.1021/acsnano.0c06946
ACS Nano XXXX, XXX, XXX−XXX