Artificial intelligence-based histopathology image analysis identifies a novel subset of endometrial cancers with distinct genomic features and unfavourable outcome

Authors

Amirali Darbandsari[1]*, Hossein Farahani[2,3]*, Matthew Wiens[2], Dawn Cochrane[4], Amy Jamieson[5], David Farnell[3,6], Pouya Ahmadvand[2], Maxwell Douglas[4], Samuel Leung[4], Purang Abolmaesumi[1], Steven JM Jones[7], Aline Talhouk[5], Stefan Kommoss[8], C Blake Gilks[3,6], David G. Huntsman[3,4]‡, Naveena Singh[3,6]‡, Jessica N. McAlpine[5]‡, Ali Bashashati[2,3]‡†

* equal contribution

‡ These authors jointly supervised this work

† Corresponding author

1. Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada.

2. School of Biomedical Engineering, University of British Columbia, Vancouver, BC, Canada.

3. Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada.

4. Department of Molecular Oncology, British Columbia Cancer Research Center, Vancouver, BC, Canada.

5. Department of Gynaecology and Obstetrics and Division of Gynaecologic Oncology, University of British Columbia, Vancouver, BC, Canada.

6. Vancouver General Hospital, Vancouver, BC, Canada.

7. Michael Smith Genome Sciences Center, British Columbia Cancer Research Center, Vancouver, BC, Canada.

8. Department of Women's Health, Tübingen University Hospital, Tübingen, Germany.

Correspondence: Ali Bashashati. Email: ali.bashashati@ubc.ca. Mailing address: 2222 Health Sciences Mall, Vancouver, BC V6T 1Z3 Canada

Abstract

Endometrial cancer (EC) has four molecular subtypes with strong prognostic value and therapeutic implications. The most common subtype (NSMP; No Specific Molecular Profile) is assigned after exclusion of the defining features of the other three molecular subtypes and includes patients with heterogeneous clinical outcomes. In this study, we employed artificial intelligence (AI)-powered histopathology image analysis to identify a novel sub-group of NSMP EC patients that had markedly inferior progression free and disease free survival in a discovery cohort of 368 patients and an independent validation cohort of 290 patients from another center. Shallow whole genome sequencing revealed a higher burden of copy number abnormalities in the identified group, compared to other NSMP EC, in our discovery and validation cohorts. Taken together, our work demonstrates the power of AI to discover new knowledge, identifying a prognostically relevant subset of EC that is unrecognizable with conventional histopathological assessment, refining image-based tumor classification.

Keywords: endometrial cancer, computational pathology, digital pathology, molecular classification, Artificial intelligence

Number of words: 3313

Number of figures: 5

Number of tables: 2

Introduction

The clinicopathological parameters used for decades to classify endometrial cancers (EC) and guide management have been sub-optimally reproducible, particularly in high-grade tumors[1,2]. Specifically, inconsistency in grade and histotype assignment has yielded inaccurate assessment of the risk of disease recurrence and death. As a result, many women affected by EC may be over-treated or are not directed to treatment that might have reduced their risk of recurrence. In 2013, the Cancer Genome Atlas (TCGA) project demonstrated that endometrial cancers could be stratified into four distinct prognostic groups using a combination of whole genome and exome sequencing, microsatellite instability (MSI) assays, and copy number analysis[3]. These subtypes were labelled according to dominant genomic abnormalities and included 'ultra-mutated' ECs harboring *POLE* mutations, 'hypermutated' identified to have microsatellite instability, copy-number low, and copy-number high endometrial cancers.

Inspired by this initial discovery, our team and a group from the Netherlands independently and concurrently developed a pragmatic, clinically applicable molecular classification system that classifies ECs into : (i) *POLE* mutant (*POLE*mut) with pathogenic mutations in the exonuclease domain of *POLE* (DNA polymerase epsilon, involved in DNA proofreading repair), (ii) mismatch repair deficient (MMRd) diagnosed based on the absence of key mismatch repair proteins on immunohistochemistry (IHC), (iii) p53 abnormal (p53abn) as assessed by IHC, and (iv) NSMP (No Specific Molecular Profile), lacking any of the defining features of the other three subtypes[4,5]. Categorization of ECs into these subtypes recapitulates the survival curves/prognostic value of the four TCGA molecular subgroups and enhances histopathological evaluation, offering an objective and reproducible classification system with strong prognostic value and therapeutic implications. In 2020, the World Health Organization (WHO) recommended integrating these key molecular features into standard pathological reporting of ECs when available[6].

*POLE*mut endometrial cancers have highly favorable outcomes with almost no deaths due to disease. While the three other molecular subtypes are associated with more variable outcomes (MMRd and NSMP are considered 'intermediate risk' and p53abn ECs have the worst prognosis), within each subtype there are clinical and prognostic outliers[7–10]. This is particularly true within the largest subtype, NSMP (representing ~50% of ECs). The majority of NSMP tumors are early stage, low grade, estrogen driven tumors likely cured by surgery alone. However, a subset of patients with NSMP EC experience a very aggressive disease course, comparable to what is observed in patients with p53abn ECs. At present, no tools exist to identify these aggressive outliers and current clinical guidelines do not stratify or direct

treatment within NSMP EC beyond using pathologic features[11,12]. Thus, for half of diagnosed endometrial cancers, i.e., NSMP EC, assumption of indolence is inappropriate and clinicians need tools for accurate risk stratification of individual patients when making treatment decisions.

With the rise of artificial intelligence (AI) in the past decade, deep learning methods (e.g., deep convolutional neural networks and their extensions) have shown impressive results in processing text and image data[13]. The paradigm shifting ability of these models to learn predictive features from raw data presents exciting opportunities with medical images, including digitized histopathology slides[14–16]. In recent years, these models have been deployed to reproduce or improve pathology diagnosis in various disease conditions (e.g.,[17–19]), explore the potential link between histopathologic features and molecular markers in different cancers including EC[20–23] , and directly link histopathology to clinical outcomes[24–27].

We hypothesized that within the NSMP molecular subtype of endometrial cancer, there is a subset of patients with aggressive disease whose tumors have histological features similar to p53abn EC and that these tumors can be identified by deep learning models applied to hematoxylin & eosin (H&E)-stained slides. As such, we have built a deep learning-based H&E image classifier that can identify distinctive histological patterns associated with p53abn and NSMP subtypes and consequently is able to identify NSMP cases that have similar histological features as p53abn EC. Our results show that these cases (referred to as *p53abn-like NSMP*) have inferior outcomes compared to the other NSMP ECs, similar to that of p53abn EC, in two independent cohorts. Furthermore, the genomic architecture of the *p53abn-like NSMP* differs from other NSMP EC, showing increased copy number abnormalities, a characteristic of p53abn EC.

Results

Patient cohort selection and description. 1,678 H&E-stained hysterectomy tissue sections from 658 patients with histologically confirmed endometrial carcinoma of NSMP or p53abn subtypes were included in this study[3–5]. Our discovery cohort included 155 whole section slides (WSI) from 146 patients from TCGA[3] and 431 WSIs (222 patients) from another center[5]. Our validation set included tissue microarray (TMA) data corresponding to 290 patients from our own center[4]. **Tables 1** and **2** show the clinicopathological features of the discovery and validation cohorts.

Histopathology-based machine learning classifier to differentiate NSMP and p53abn ECs. Fig. 1 depicts our AI-based histopathology image analysis pipeline. A subset of 27 whole section H&E slides from the TCGA cohort were annotated by a board-certified pathologist (DF) using a custom in-house histopathology

slide viewer (*cPathPortal*) to identify areas containing tumor and stromal cells. A deep convolutional neural network (CNN)-based classifier was then trained to acquire pseudo-tumor and benign annotations for the remaining slides in the discovery cohort (see the AI tumor-normal classifier and automatic annotation section of the Methods). The identified tumor regions were then divided into 512x512 pixel patches at 20x objective magnification. The number of extracted patches from each subtype and performance measure for the tumor stroma classifier can be found in **Supplementary Tables 1 and 2**. To address variability in slide staining due to differences in staining protocols across different centres, and inter-patient variability, we utilized the Vahadane color normalization technique[28]. We then trained a VarMIL model[29] based on multiple instance learning (MIL) to differentiate H&E image patches associated with p53abn and NSMP ECs (see the Deep learning model for tumor subtype classification section of the Methods).

In a "group 10-fold" cross-validation strategy, the patients in our discovery cohort were divided into 10 groups and in various combinations, 60% were used for training, 20% for validation, and 20% for testing; resulting in 10 different binary p53abn vs. NSMP classifiers. These 10 classifiers were then used to label the cases as p53abn or NSMP and their consensus was used to come up with a label for a given case. For patients with multiple slides, to prevent data leakage between training, validation, and test sets, we assigned slides from each patient to only one of these sets.

Fig. 2A and **Supplementary Table 3** show the receiver operating characteristics (ROC) and precision/recall curves as well as performance metrics of the resulting classifiers for the discovery and validation sets, respectively. These results suggest that our p53abn vs. NSMP classifier achieves more than 89.4% and 79.8% mean balanced accuracy (across the 10 classifiers) in both the discovery and validation sets, respectively (for details see **Supplementary Table 4 and Extended Data Fig. 1**).

Identification of a subset of NSMP ECs with inferior survival. Our proposed ML-based models classified 17.65% and 20% of NSMPs as p53abn for the discovery and validation cohorts, respectively (**Supplementary Table 5** and the Identification of *p53abn-like NSMP*s section of the Methods). These cases (referred to as *p53abn-like NSMP* group) show p53abn histological features in the assessment of H&E images even though immunohistochemistry did not show mutant-pattern p53 expression and these were therefore classified as NSMP by the molecular classifier. We hypothesized that such cases may in fact exhibit similar clinical behavior as p53abn ECs.

Fig. 2B,C show the progression free survival (PFS) and disease specific survival (DSS) of the discovery and validation sets (Survival analysis section of the Methods). Compared to the rest of the NSMP cases,

*p53abn-like NSMP*s had markedly inferior PFS (10-year PFS 55.7% vs. 89.6% (p < 2.7e-7)) and DSS (10-year DSS 62.6% vs. 93.7% (p < 1.8e-7)) in our discovery cohort. These findings were confirmed in the validation cohort, with 20% of the 195 patients categorized as p53abn-like tumors, showing 10-year PFS of 65.4% vs. 91.2% (p < 1.1e-4) and DSS of 58.3% vs. 84.3% (p < 5.3e-5). Additionally, comparison of the PFS and DSS between *p53abn-like NSMP* and p53abn ECs revealed a trend, though not statistically significant, in which *p53abn-like NSMP*s had better outcome compared to p53abn ECs in both the discovery and validation cohorts (**Extended Data Fig. 2A,B**).

Of note, our model identified a subset of p53abn ECs (representing 20%; referred to as *NSMP-like p53abn*) with resemblance to NSMP as assessed by H&E staining. While we observed marginally superior disease-specific survival in the identified cases compared to the rest of the p53abn group both in the discovery and validation cohorts, progression free survival was not significantly different between the groups (**Extended Data Fig. 3A,B**).

Comparison of NSMP and *p53abn-like NSMP*. To further investigate the differences between NSMP and *p53abn-like NSMP* cases, we compared various clinical, pathological, and molecular variables (**Supplementary Tables 6 and 7B).** Our analysis showed an enrichment of *p53abn-like NSMP* cases with higher grade and higher stage tumors (p < 1.4e-25; p < 2.4e-4, respectively). In a multi-variate Cox regression analysis, the prognostic significance of the *p53abn-like NSMP* group remained significant in the presence of grade and stage (p = 0.02 and Hazard Ratio = 2.44; **Supplementary Table 8**). Furthermore, **Fig. 3A** shows an enrichment for estrogen receptor (ER) and progesterone receptor (PR) positive cases in the *p53abn-like NSMP*s in the subset of the cohort that the status these markers were available (p < 5.2e-3 and p < 2.3e-4, respectively).

Independent pathology review of selected NSMP cases. Two expert gynepathologists (NS, CBG) independently reviewed whole section slides of a subset of NSMP cases including the *p53abn-like NSMP* subtype. They specifically assessed whether tumors showed nuclear features that have been previously described as being associated with *TP53* mutation/mutant pattern p53 expression in endometrial carcinoma[30]. The *p53abn-like NSMP* cases were enriched with tumors showing increased nuclear atypia, as assessed by altered chromatin pattern, nucleolar features, pleomorphism, atypical mitoses, or giant tumor cells (p < 0.00005 for both reviewers).

Genomic characterization of *p53abn-like NSMP* cases. We next sought to investigate the molecular profiles of *p53abn-like NSMP* cases in our validation set for which we had access to tissue material. Targeted sequencing of exonic regions in a number of genes (more details in the Targeted point mutation

profiling section of the Methods) revealed enrichment of *p53abn-like NSMP* cases with *TP53* mutations and enrichment of NSMP cases with *CTNNB1* mutations (Fisher's exact test p-values = 3.14e-4 and 0.01, respectively; **Fig. 3A**). More specifically, we identified eight *p53abn-like NSMP* tumors that had normal p53 IHC results (hence classified as NSMP by ProMisE classifier) but in fact harbored *TP53* mutations by sequencing. These cases are examples of the well-known phenomenon of normal p53 protein levels despite there being a pathogenic mutation, which occurs in <5% of cases. However, even after removing these eight *TP53* mutant cases, the worse prognosis of *p53abn-like NSMP* tumors persisted (**Fig. 3B**). Our ML model, therefore, identifies tumors with false negative immunostaining for p53 ,i.e., they lack mutant pattern protein expression despite having a *TP53* mutation, but also identifies a subset of NSMP cases with features of p53abn morphology by H&E but neither mutation pattern immunostaining nor a mutation in sequencing *TP53*, and these have inferior survival compared to the rest of the NSMP cases. Exonic point mutation data for the TCGA subset of the discovery cohort were available and suggested a lack of enrichment of the *p53abn-like NSMP* group with specific gene mutations including *CTNNB1* (**Fig. 3C**).

We next selected representative samples of NSMP, p53abn, and *p53abn-like NSMP* cases in our validation cohort and performed shallow whole genome sequencing (sWGS). Overall, copy number profile analysis of these cases revealed that *p53abn-like NSMP* cases harbor a higher fraction of altered genome compared to NSMP cases but still lower than what we observe in p53abn cases (**Fig. 4A**; $p < 0.035$). These findings were further validated in the TCGA cohort (**Fig. 4B**; $p < 5.46e-5$).

We next investigated the gene expression profiles associated with the *p53abn-like NSMP*, NSMP, and p53abn tumors within the TCGA cohort. Unsupervised clustering of patients based on gene expression profiles of their tumors showed that eleven of the 21 *p53abn-like NSMPs* had similar expression profiles to p53abn tumors, while the remaining 10 cases clustered together with the NSMP group (**Fig. 4C**). While p53abn and NSMP groups were clustered separately, unsupervised analysis of the gene expression profiles did not reveal any differences between *p53abn-like NSMP* group and other subtypes, i.e., they did not have a unique gene expression profile but instead clustered with one of the known molecular subtypes. We then performed pairwise differential expression analysis and pathway analysis, separately comparing *p53abn-like NSMP* and p53abn groups against NSMP cases. These results suggested the upregulation of PI3k-Akt, Wnt, and Cadherin signaling pathways both in *p53abn-like NSMP* and p53abn groups (compared to NSMP). Interestingly, while these pathways were up-regulated in both groups, we found little to no overlap between the specific down- and up-regulated genes in the *p53abn-like NSMP* and p53abn groups (compared to NSMP) suggesting that the molecular mechanisms associated with

p53abn and p53abn-like tumors might be different even though p53abn and *p53abn-like NSMP* groups had similar histopathological profiles as assessed based on H&E slides (the Gene expression analysis section of the Methods).

Discussion

Although many patients with endometrial carcinoma may be cured by surgery alone, about 1 in 5 patients have more aggressive disease and/or have the disease spread beyond the uterus at the time of diagnosis. Identifying these at-risk individuals remains a challenge, with current tools lacking precision. Molecular classification offers an objective and reproducible classification system that has strong prognostic value; improving the ability to discriminate outcomes compared to conventional pathology-based risk stratification criteria. However, it has become apparent that within molecular subtypes and most profoundly within NSMP ECs, there are clinical outcome outliers. The current study addresses this diversity by employing AI-powered histopathology image analysis, in an attempt to identify clinical outcome outliers within the most common molecular subtype of endometrial cancer (**Fig. 5**). Our results have several clinical and biological implications.

To be clear, for some molecular subtypes, such as *POLE*mut endometrial cancers with almost uniformly favourable outcomes, no further stratification, at least within Stage I-II disease (encompassing >90% of *POLE*mut ECs), is needed. Multiple studies, as well as meta-analyses[31], have shown that in patients with *POLE*mut endometrial cancers, additional pathological or molecular features are not associated with outcomes, i.e., are not prognostic, as *POLE* is the overriding feature that determines survival. However, for NSMP endometrial cancers, additional stratification tools are greatly needed. Designation of NSMP is the last step in molecular classification, only defined by what molecular features it does ***not*** have; that is without pathogenic *POLE* mutations, without mismatch repair deficiency or p53 abnormalities as assessed by IHC. This leaves a large group of pathologically and molecularly diverse tumors with markedly varied clinical outcomes.

Our AI-based histomorphological image analysis model identified a previously unrecognized subset of NSMP endometrial cancers with inferior survival. This subset of patients encompassed approximately 20% of NSMP tumors which are the most common molecular subtype, representing half of endometrial cancers diagnosed in the general population, and thus account for 10% of all ECs. Our results suggest that clinicopathological, IHC, gene expression profiles, or NGS molecular markers (except for copy number burden to some extent) may not be able to identify these p53abn-like outliers. The AI classifier was able to identify those tumors with *TP53* mutations (but normal p53 immunostaining), a result we view as

encouraging, in that these are “false negative” cases using the IHC classification and more appropriately assigned as p53abn, but even when these were removed from consideration AI imaging discerned other patients with NSMP EC where no molecular or pathological features would have identified them as having inferior outcomes. AI applied to histomorphological images of routinely generated H&E slides appears to enable a more encompassing and easily implementable stratification of NSMP tumors and provides greater value than any single or combined pathological/molecular profile could achieve.

Molecular characterization of the identified subtype using sWGS suggests that these cases harbor an unstable genome with a higher fraction of altered genome, similar to the p53abn group but with lesser degree of instability. In spite of the fact that similar gene expression pathways were implicated in both groups and H&E images of both groups as assessed by AI had resemblances, expression data analysis revealed minimal overlap between the differentially expressed genes in both p53abn and p53abn-like EC compared to NSMP cases. This suggests that they may have different etiologies and warrants further biological interrogation of these groups in future studies.

Certainly, others have attempted to refine stratification within early-stage endometrial cancers, including within the molecularly defined NSMP subset. PORTEC4a used a combination of pathologic and molecular features (MMRd, L1CAM overexpression*, POLE*, *CTNNB1* status) to identify low, intermediate, and high risk individuals assigned to favourable, intermediate, and unfavourable risk groups which then determined observations vs. treatment[32]. TAPER/EN.10 also stratifies early-stage NSMP tumors by pathological (e.g., histotype, grade, LVI status) and molecular features (*TP53,* ER status) to identify those individuals appropriate for de-escalated therapy[33]. In retrospective series, key parameters of ER and grade have been suggested to discern outcomes within NSMP. However, even in-depth profiling of apparent low risk ECs has failed to find pathogenomic features that would discern individuals who develop recurrence from other apparent indolent tumors[34]. Stasenko et al.[34] assessed a series of 486 cases of ‘ultra-low risk’ endometrial cancers defined as stage 1A with no myoinvasion, no LVI, grade 1 of which 2.9% developed recurrence with no identifiable associated clinical, pathological or molecular features[34]. Current treatment guidelines, even where molecular features are incorporated, offer little in terms of directing management within NSMP endometrial cancers beyond consideration of pathological features, leaving clinicians to struggle with optimal management[12]. A more comprehensive stratification tool within NSMP endometrial cancers would be of tremendous value, and AI discernment from histopathological images as a tool that can be readily applied to H&E slides that are routinely generated as part of the practice is appealing.

Our proposed AI model also identified a subset of p53abn ECs with marginally superior DSS and resemblance to NSMP (*NSMP-like p53abn*) as assessed by H&E staining. Further investigation of the identified groups and deep molecular and omics characterization of this subset of p53abn ECs may in fact aid us in refining this subtype and identifying a subset of p53abn cases with statistically superior outcomes.

This study is the first to consider the application of AI in refining endometrial cancer molecular subtypes. In general, such studies to generate new knowledge using AI in histopathology are extremely sparse as a majority of the effort has focused on recapitulating the existing body of knowledge (e.g., to diagnose cancer, to identify histological subtypes, to identify known molecular subtypes). This study moves beyond the mainstream AI applications and more specifically identifies a new prognostic subset of EC, which also then enables us to direct efforts to understand the biological mechanisms of this newly identified subset. This could present an exciting opportunity to utilize the power of AI to inform clinical trials and deep biological interrogation by adding more precision in patient stratification and selection.

AI histopathologic imaging-based application within NSMP offers the chance to discern outcomes within the largest endometrial cancer molecular subtype. It can be easily added to clinical algorithms after performing hysterectomy, identifying some patients (*p53abn-like NSMP*) as candidates for treatment analogous to what is given in p53abn tumors. Furthermore, the proposed AI model can potentially have greater impact on patient management and equitable cancer care if confirmed in diagnostic biopsies. If, from diagnostic office biopsy or surgical curettage endometrial cancers could be classified as NSMP tumors and then AI stratification applied, we would have the opportunity to guide what surgery is warranted, potentially directing individuals at very low risk of metastases to simple hysterectomy in the community and more aggressive *p53abn-like NSMP* to cancer centers for lymph node assessment, omental sampling and directed biopsies given a higher likelihood of upstaging.

References

1. Gilks, C. B., Oliva, E. & Soslow, R. A. Poor interobserver reproducibility in the diagnosis of high-grade endometrial carcinoma. *The American journal of surgical pathology* **37**, 874–881 (2013).
2. Hoang, L. N. *et al.* Histotype-genotype correlation in 36 high-grade endometrial carcinomas. *The American journal of surgical pathology* **37**, 1421–1432 (2013).
3. Levine, D. A. *et al.* Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67–73 (2013).
4. Talhouk, A. *et al.* Confirmation of ProMisE: A simple, genomics-based clinical classifier for endometrial cancer. *Cancer* **123**, 802–813 (2017).
5. Kommoss, S. *et al.* Final validation of the ProMisE molecular classifier for endometrial carcinoma in a large population-based case series. *Annals of Oncology* **29**, 1180–1188 (2018).
6. Editorial Board, W. C. of T. *WHO Classification of Tumours Female Genital Tumours*. vol. Volume 4 (International Agency for Research on Cancer, 2020).
7. Kasius, J. C. *et al.* Risk Stratification of Endometrial Cancer Patients: FIGO Stage, Biomarkers and Molecular Classification. *Cancers* **13**, 5848 (2021).
8. Thompson, E. *et al.* Further stratification of no specific molecular profile (NSMP/P53WT) endometrial carcinomas to refine prognosis and identify therapeutic opportunities. *International Journal of Gynecologic Cancer* **31**, A17–A17 (2021).
9. Leo, A. D. *et al.* ARID1A and CTNNB1/β-Catenin Molecular Status Affects the Clinicopathologic Features and Prognosis of Endometrial Carcinoma: Implications for an Improved Surrogate Molecular Classification. *Cancers* **13**, 950 (2021).
10. Kolehmainen, A. *et al.* Clinical factors as prognostic variables among molecular subgroups of endometrial cancer. *PLoS One* **15**, e0242733 (2020).
11. Prakasan, A. M. *et al.* The Pattern of Recurrence in Carcinoma Endometrium. *Indian Journal of Gynecologic Oncology* **20**, 1–7 (2022).
12. National Comprehensive Cancer Network (NCCN). Uterine Neoplasms NCCN Guidelines Version 4.2021. (2021).
13. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
14. Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V. & Madabhushi, A. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nat Rev Clin Oncol* **16**, 703–715 (2019).
15. Srinidhi, C. L., Ciga, O. & Martel, A. L. Deep neural network models for computational histopathology: A survey. (2019).
16. Shmatko, A., Ghaffari Laleh, N., Gerstung, M. & Kather, J. N. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. *Nat Cancer* **3**, 1026–1038 (2022).
17. Bejnordi, B. E. *et al.* Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women With Breast Cancer. *JAMA* **318**, 2199–2210 (2017).
18. Bulten, W. *et al.* Artificial intelligence for diagnosis and Gleason grading of prostate cancer: the PANDA challenge. *Nat Med* **28**, 154–163 (2022).
19. Campanella, G. *et al.* Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* **25**, 1301–1309 (2019).
20. Coudray, N. *et al.* Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nature Medicine* 1 (2018) doi:10.1038/s41591-018-0177-5.
21. Fu, Y. *et al.* Pan-cancer computational histopathology reveals mutations, tumor composition and prognosis. *bioRxiv* 813543 (2019) doi:10.1101/813543.
22. Kather, J. N. *et al.* Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. *Nat Med* **25**, 1054–1056 (2019).

23. Hong, R., Liu, W., DeLair, D., Razavian, N. & Fenyö, D. Predicting endometrial cancer subtypes and molecular features from histopathology images using multi-resolution deep learning models. *CR Med* **2**, (2021).
24. Wetstein, S. C. *et al.* Deep learning-based breast cancer grading and survival analysis on whole-slide histopathology images. *Sci Rep* **12**, 15102 (2022).
25. Lee, Y. *et al.* Derivation of prognostic contextual histopathological features from whole-slide images of tumours via graph deep learning. *Nat. Biomed. Eng* 1–15 (2022) doi:10.1038/s41551-022-00923-0.
26. Zadeh Shirazi, A. *et al.* A deep convolutional neural network for segmentation of whole-slide pathology images identifies novel tumour cell-perivascular niche interactions that are associated with poor survival in glioblastoma. *Br J Cancer* **125**, 337–350 (2021).
27. Jiang, S., Zanazzi, G. J. & Hassanpour, S. Predicting prognosis and IDH mutation status for patients with lower-grade gliomas using whole slide images. *Sci Rep* **11**, 16849 (2021).
28. Vahadane, A. *et al.* Structure-Preserving Color Normalization and Sparse Stain Separation for Histological Images. *IEEE Transactions on Medical Imaging* **35**, 1962–1971 (2016).
29. Schirris, Y., Gavves, E., Nederlof, I., Horlings, H. M. & Teuwen, J. DeepSMILE: Contrastive self-supervised pre-training benefits MSI and HRD classification directly from H&E whole-slide images in colorectal and breast cancer. *Medical Image Analysis* **79**, 102464 (2022).
30. Kang, E. Y. *et al.* Selection of endometrial carcinomas for p53 immunohistochemistry based on nuclear features. *The Journal of Pathology: Clinical Research* **8**, 19–32 (2022).
31. McAlpine, J. N. *et al.* Evaluation of treatment effects in patients with endometrial cancer and POLE mutations: An individual patient data meta-analysis. *Cancer* **127**, 2409–2422 (2021).
32. Wortman, B. *et al.* Molecular-integrated risk profile to determine adjuvant radiotherapy in endometrial cancer: evaluation of the pilot phase of the PORTEC-4a trial. *Gynecologic oncology* **151**, 69–75 (2018).
33. ClinicalTrials.gov identifier (NCT number): NCT04705649. Tailored Adjuvant Therapy in POLE-mutated and p53-wildtype Early Stage Endometrial Cancer (TAPER).
34. Stasenko, M. *et al.* Clinical patterns and genomic profiling of recurrent 'ultra-low risk'endometrial cancer. *International Journal of Gynecologic Cancer* **30**, (2020).

Ethics statement. The Declaration of Helsinki and the International Ethical Guidelines for Biomedical Research Involving Human Subjects were strictly adhered throughout the course of this study. All study protocols have been approved by the University of British Columbia/BC Cancer Research Ethics Board.

List of Figures:

Fig. 1: Workflow of the AI-based histopathology image analysis. First, the quality control framework, HistoQC[57], generates a mask that comprises tissue regions exclusively and removes artifacts. Then, an AI model to identify tumor regions within histopathology slides is trained. Next, images are tessellated into small patches and normalized to remove color variations. The normalized patches are fed to a deep learning model to derive patch-level representations. Finally, a model based on multiple instance learning (VarMIL) was utilized to predict the patient subtype.

Fig. 2: Performance statistics and Kaplan Meier (KM) survival curves for AI-identified EC subtypes. **(A)** AUROC and precision-recall plots of average of 10 splits for p53abn vs. NSMP classifier for discovery and validation sets, **(B)** KM curves associated with PFS and DSS for the discovery set, **(C)** KM curves associated with PFS and DSS (where available) in the validation set.

Fig. 3: (A) Clinicopathological features and point mutation data for the validation cohort, **(B)** KM curves associated with PFS and DSS for the validation cohort after removal of the samples with *TP53* mutations, **(C)** Clinicopathological features and point mutation data for the TCGA cohort.

Fig. 4: Molecular profiling of *p53abn-like NSMP* cases. Boxplots of copy number burden (i.e., fraction genome altered) in NSMP, *p53abn-like NSMP*, and p53abn cases in the **(A)** validation (6 NSMP, *7 p53abn-like NSMP*, 5 p53abn) and **(B)** TCGA (69 NSMP, *21 p53abn-like NSMP*, 56 p53abn) cohorts. The box plot indicates the median, with the box representing the first to third quartile and the whiskers extending 1.5 times the interquartile range. **(C)** Gene expression profiles associated with the *p53abn-like NSMP* (n=21), NSMP (n=69), and p53abn (n=56) tumors in the TCGA cohort.

Fig. 5: The refined classification scheme that leverages AI screening as a supplementary stratification mechanism within the NSMP molecular subtype.

List of Tables:

Table 1: Clinicopathologic features of the discovery set.

Table 2: Clinicopathologic features of the validation set.

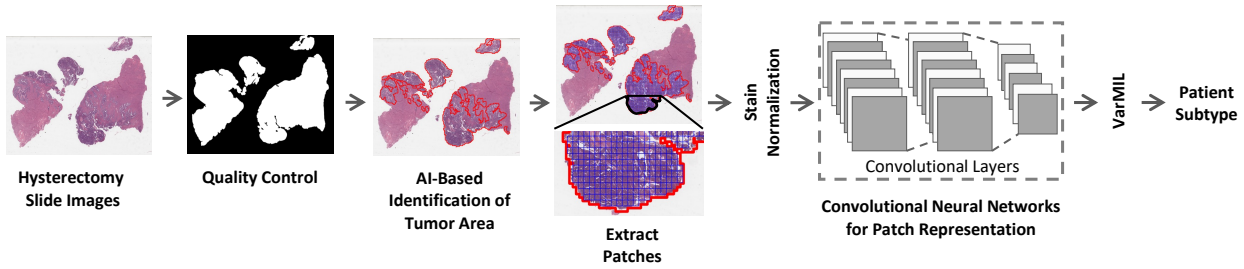

Fig. 1: Workflow of the AI-based histopathology image analysis. First, the quality control framework, HistoQC[57], generates a mask that comprises tissue regions exclusively and removes artifacts. Then, an AI model to identify tumor regions within histopathology slides is trained. Next, images are tessellated into small patches and normalized to remove color variations. The normalized patches are fed to a deep learning model to derive patch-level representations. Finally, a model based on multiple instance learning (VarMIL) was utilized to predict the patient subtype.

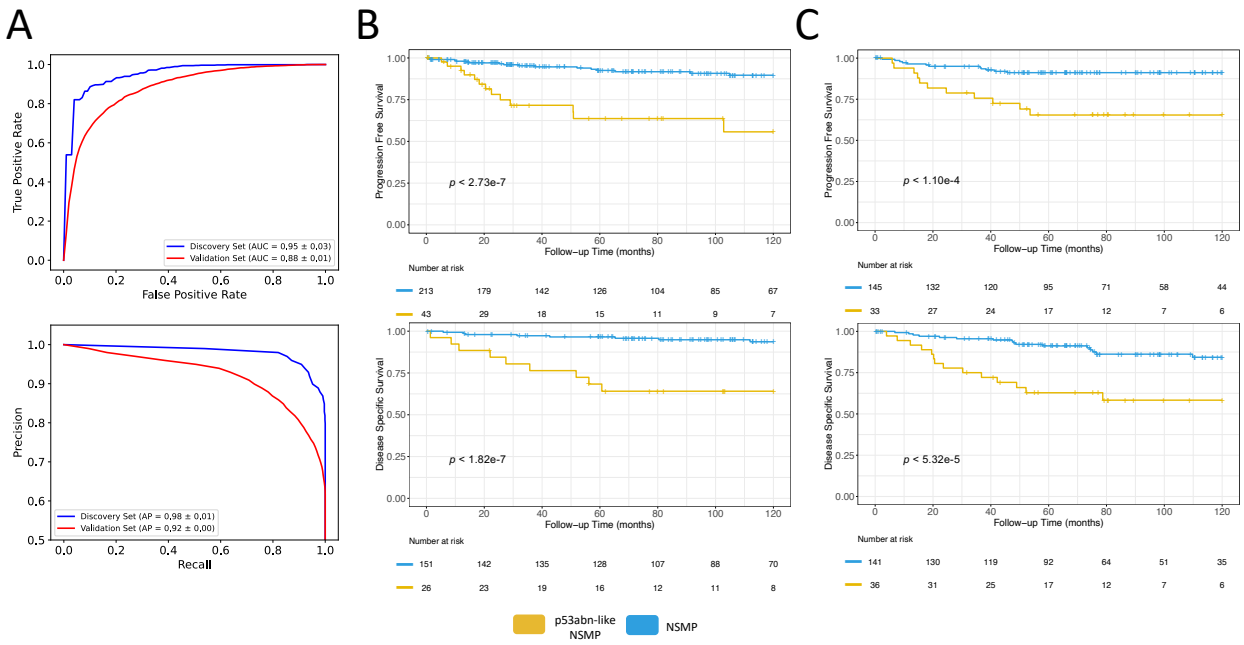

Fig. 2: Performance statistics and Kaplan Meier (KM) survival curves for AI-identified EC subtypes. **(A)** AUROC and precision-recall plots of average of 10 splits for p53abn vs. NSMP classifier for discovery and validation sets, **(B)** KM curves associated with PFS and DSS for the discovery set, **(C)** KM curves associated with PFS and DSS (where available) in the validation set.

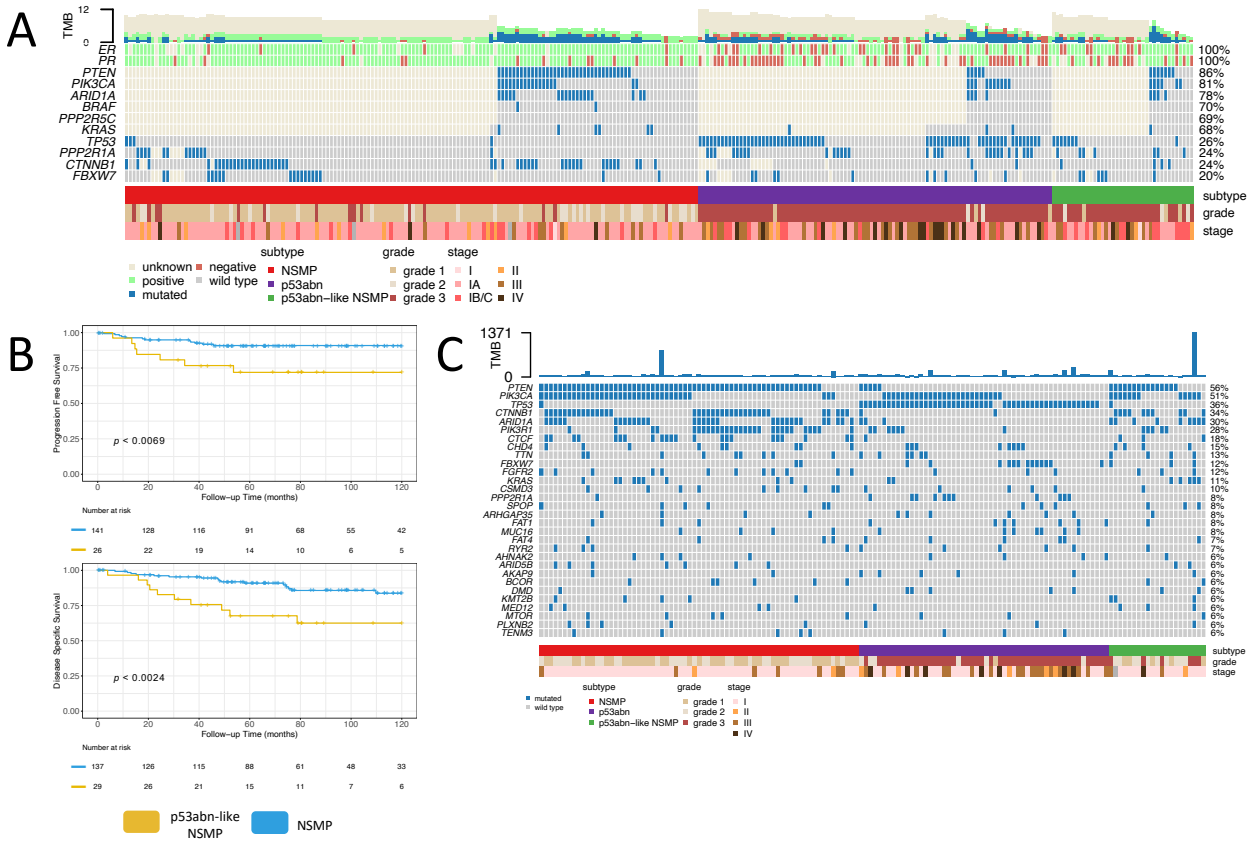

Fig. 3: (A) Clinicopathological features and point mutation data for the validation cohort, **(B)** KM curves associated with PFS and DSS for the validation cohort after removal of the samples with *TP53* mutations, **(C)** Clinicopathological features and point mutation data for the TCGA cohort.

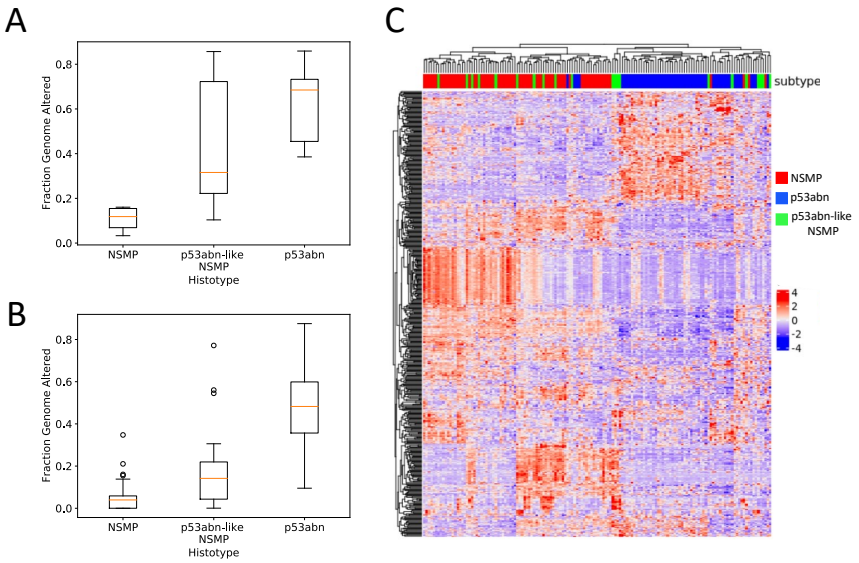

Fig. 4: Molecular profiling of *p53abn-like NSMP* cases. Boxplots of copy number burden (i.e., fraction genome altered) in NSMP, *p53abn-like NSMP*, and p53abn cases in the **(A)** validation (6 NSMP, *7 p53abn-like NSMP*, 5 p53abn) and **(B)** TCGA (69 NSMP, *21 p53abn-like NSMP*, 56 p53abn) cohorts. The box plot

indicates the median, with the box representing the first to third quartile and the whiskers extending 1.5 times the interquartile range. **(C)** Gene expression profiles associated with the *p53abn-like NSMP* (n=21), NSMP (n=69), and p53abn (n=56) tumors in the TCGA cohort.

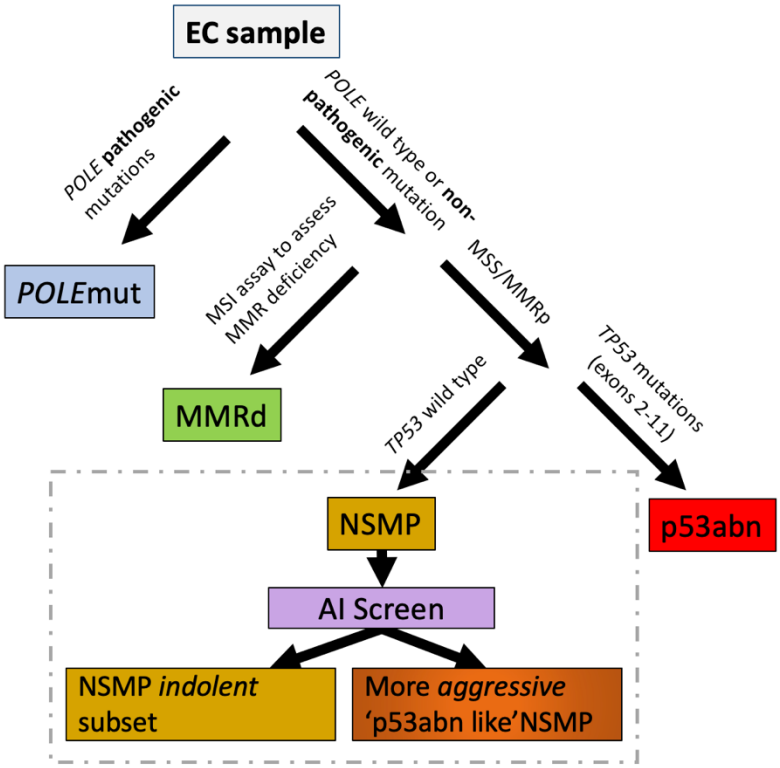

Fig. 5: The refined classification scheme that leverages AI screening as a supplementary stratification mechanism within the NSMP molecular subtype.

Table 1: Clinicopathologic features of the discovery set.

Variable	Total	NSMP	p53abn
Total	363	268 (73.83%)	95 (26.17%)
Age at diagnosis			
<60 yrs	121 (33.33%)	110 (41.04%)	11 (11.58%)
≥60 yrs	242 (66.67%)	158 (58.96%)	84 (88.42%)
Histotype			
Endometrioid	288 (79.34%)	262 (97.76%)	26 (27.37%)
Non-endometrioid	75 (20.66%)	6 (2.24%)	69 (72.63%)
Tumor grade			
Low grade (G1–2)	258 (71.07%)	246 (91.79%)	12 (12.63%)
High grade (G3)	105 (28.93%)	22 (8.21%)	83 (87.37%)
FIGO stage			
I-II	291 (80.17%)	239 (89.18%)	52 (54.74%)
III-IV	71 (19.56%)	28 (10.45%)	43 (45.26%)
Unknown	1 (0.28%)	1 (0.37%)	0

Table 2: Clinicopathologic features of the validation set.

Variable	Total	NSMP	p53abn
Total	288	193 (67.01%)	95 (32.99%)
Age at diagnosis			
<60 yrs	81 (28.13%)	72 (37.70%)	9 (9.47%)
≥60 yrs	205 (71.18%)	119 (62.30%)	86 (90.53%)
Unknown	2 (0.69%)	2 (1.04%)	0
Histotype			
Endometrioid	195 (67.71%)	172 (89.12%)	23 (24.21%)
Non-endometrioid	91 (31.60%)	19 (9.84%)	72 (75.79%)
Unknown	2 (0.69%)	2 (1.04%)	0
Tumor grade			
Low grade (G1–2)	151 (52.43%)	146 (75.65%)	5 (5.26%)
High grade (G3)	137 (47.57%)	47 (24.35%)	90 (94.74%)
FIGO stage			
I-II	216 (75.00%)	166 (86.01%)	50 (52.63%)
III-IV	69 (23.96%)	24 (12.44%)	45 (47.37%)
Unknown	3 (1.04%)	3 (1.55%)	0