

INTEGRATED IMAGE-BASED IMMUNE AND MOLECULAR PROFILING REFINES PROGNOSTICATION IN EARLY-STAGE ENDOMETRIAL CANCER

Running head: Integrated immune-molecular profiling in endometrial cancer

Authors

Nanda Horeweg^{1*}, Marco de Bruyn^{2*}, Remi A. Nout^{1~}, Ellen Stelloo³, Katarzyna Kedziersza⁴, Alicia León-Castillo³, Annechien Plat², Kirsten D Mertz⁵, Michelle Osse³, Ina M. Jürgenliemk-Schulz⁶, Ludy C.H.W. Lutgens⁷, Jan J. Jobsen⁸, Elzbieta M. van der Steen-Banasik⁹, Vincent T.H.B.M. Smit³, Carien L. Creutzberg¹, Tjalling Bosse³, Hans W. Nijman², Viktor H Koelzer^{10,11†}, David N Church^{3,12,13‡,‡}

Affiliations

¹ Department of Radiation Oncology, Leiden University Medical Center, Leiden, The Netherlands.

² Department of Gynaecologic Oncology, University Medical Center Groningen, Groningen, The Netherlands

³ Department of Pathology, Leiden University Medical Center, Leiden, The Netherlands.

⁴ Wellcome Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford, OX3 7BN, UK

⁵ Cantonal Hospital Baselland, Institute of Pathology, Liestal, Switzerland

⁶ Department of Radiation Oncology, University Medical Center Utrecht, Utrecht, The Netherlands

⁷ Maastricht Radiation Oncology Clinic, Maastricht, The Netherlands.

⁸ Department of Radiotherapy, Medisch Spectrum Twente, Enschede, The Netherlands.

⁹ Radiotherapiegroep, Arnhem, The Netherlands.

¹⁰ Department of Pathology and Molecular Pathology, University of Zurich, Switzerland

¹¹ Department of Oncology and Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom

¹² Oxford Cancer Centre, Churchill Hospital, Oxford University Hospitals Foundation NHS Trust, Oxford, UK

¹³ Oxford NIHR Comprehensive Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, Oxford, UK

* Denotes equal contribution

~ Currently employed at department of Radiation Oncology, Erasmus MC- Cancer Institute, Rotterdam, The Netherlands

† Denotes joint senior author

‡ Denotes corresponding author

Tel: +44 (0)1865 287500; Fax: +44 (0)1865 287501; email: dchurch@well.ox.ac.uk

Word count

3642

RESEARCH SUPPORT

The PORTEC-1 and PORTEC-2 trials were supported by grants from the Dutch Cancer Society (CKTO 90-01 and CKTO 2001-04 respectively). This study was supported by a translational research project grant from the Dutch Cancer Society (UL2012-5719), the Oxford NIHR Comprehensive Biomedical Research Centre (BRC), and core funding to the Wellcome Trust Centre for Human Genetics from the Wellcome Trust (090532/Z/09/Z). KK is supported by a Genomic Medicine and Statistics Studentship from the Wellcome Trust. TB is supported by a Dutch Cancer Society Young Investigator Award. VHK is supported by the Promedica Foundation (F-87701-41-01). DNC is funded by a Cancer Research UK Advanced Clinician Scientist Fellowship and was previously supported by a Health Foundation/Academy of Medical Sciences Clinician Scientist Fellowship. The cost of open access publication were provided by core funding to the Wellcome Centre for Human Genetics from the Wellcome Trust (203141/Z/16/Z). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, the Department of Health or the Wellcome Trust.

PREVIOUS PRESENTATION

This study has not previously been presented elsewhere

DISCLAIMERS

None

ABSTRACT

Current optimum risk stratification in early-stage endometrial cancer (EC) combines clinicopathological factors and the molecular EC classification defined by The Cancer Genome Atlas (TCGA). It is unclear whether analysis of intratumoral immune infiltrate improves this.

We developed a machine-learning image-based algorithm to quantify density of CD8⁺ and CD103⁺ immune cells in tumor epithelium and stroma in 695 stage I endometrioid ECs from the PORTEC-1&-2 trials. The relationship between immune cell density and clinicopathological/molecular factors was analyzed by hierarchical clustering and multiple regression. The prognostic value of immune infiltrate by cell type and location was analyzed by univariable and multivariable Cox regression, incorporating the molecular EC classification.

Tumor-infiltrating immune cell density varied substantially between cases, and more modestly by immune cell type and location. Clustering revealed three groups with high, intermediate and low densities, with highly significant variation in the proportion of molecular EC subgroups between them ($P<0.0001$). Univariable analysis revealed intraepithelial CD8⁺ cell density as the strongest predictor of EC recurrence; multivariable analysis confirmed this was independent of pathological factors and molecular subgroup (HR=0.88 per two-fold increase; 95% CI=0.81–0.95; $P=1.8\times10^{-3}$, or HR=0.50, 95% CI=0.33–0.78, for cases at the 90th vs 10th percentile of density). Exploratory analysis suggested this association was not uniform across molecular subgroups, but strongest in tumors with mutant p53 (HR=0.84; 95% CI=0.73–0.97; $P=0.017$) and absent in DNA mismatch repair deficient cancers (HR=0.98; 95% CI=0.84–1.16; $P=0.85$).

Quantification of intraepithelial CD8⁺ cells improves prognostication beyond the current integrated molecular classification in early-stage EC.

STATEMENT OF TRANSLATIONAL RELEVANCE

Endometrial cancer (EC) is the most common gynecological malignancy in developed nations. Approximately three-quarters of affected women are diagnosed at stage I with disease confined to the uterus. Their treatment after hysterectomy is guided by the predicted likelihood of recurrence and varies from surveillance, to vault brachytherapy, external beam radiotherapy and adjuvant chemotherapy. The current optimum for estimating this combines clinicopathological factors with the molecular EC classifier defined by The Cancer Genome Atlas (TCGA). However, this is imperfect, and over- and undertreatment remain common. We investigated whether image-based quantification of intratumoral T cell density could improve this. As anticipated, T cell density varied by EC molecular subgroup, however, intraepithelial CD8⁺ cell density was strongly prognostic independent of clinicopathological and molecular factors. A combination of pathological, molecular and immune factors outperformed the current state-of-the-art and holds potential to improve prognostication in this common disease.

INTRODUCTION

Endometrial cancer (EC) is the most common gynecological malignancy in developed countries [1, 2]. Most cases are detected at early stage with disease confined to the uterus (FIGO stage I), and are managed by curative-intent surgical resection. Adjuvant external beam radiotherapy (EBRT) or vaginal vault brachytherapy (VBT) reduce the risk of pelvic recurrence, but at the expense of added toxicities [3-5]. Adjuvant chemoradiotherapy has recently been shown to reduce recurrence and improve survival in high risk cases [6]. The definition of risk groups, and thus the selection of which women should receive adjuvant local or systemic therapies has traditionally been based on clinical and pathological factors, including patient age, stage, histological type and grade, lymphovascular space invasion (LVSI), and myometrial invasion [7-9]. However, the usefulness of several of these factors in early-stage endometrioid EC is limited by their relatively modest effect size [9-11], resulting in considerable over- and under-treatment. This shortcoming has motivated intense investigation for novel EC subgroups or tumor-associated biomarkers of prognostic value, using established and emerging technologies [12-15]. Arguably the most impactful of these studies is that by The Cancer Genome Atlas [14]. This used whole exome sequencing to define four EC subgroups with differing biology and clinical outcome: (i) POLE ultramutated (POLEmut), defined by pathogenic mutations within the DNA polymerase epsilon catalytic subunit (*POLE*) exonuclease domain; (ii) DNA mismatch repair deficient (MMRd); (iii) TP53 mutant/somatic copy number alteration (SCNA) high, and; (iv) an SCNA low group lacking these other genomic alterations, and often referred to as no specific molecular profile (NSMP). Approximation of this molecular classification using surrogate markers is feasible in clinical practice [16], and improves upon prognostication provided by clinicopathological

factors [11]. Prospective evaluation of this classification, together with other promising molecular markers such as *CTNNB1* mutation [10, 17], is underway in both trial [18] and non-trial settings. Another important area of EC biomarker research pertains to the anti-tumor T-cell response, the prognostic value of which has been shown in several studies [19, 20]. While the potential clinical utility of both the molecular EC classification and quantification of the antitumor immune response appear considerable, the variable prognosis of molecular EC subgroups may be explained by a marked variation in intratumoral T-cell infiltrate between them [21, 22], and it is unclear whether these genomic and immune biomarkers confer independent prognostic value. A recent study of an EC cohort of mixed stages and histotypes found that while application of the molecular classification improved upon the prognostication provided by clinicopathological variables, the additional analysis of T-cell infiltrate conferred no further benefit [23]. We examined this in a more homogenous population of stage I endometrioid ECs from two large randomized clinical trials.

METHODS

Design and study population

Details of the PORTEC-1 and PORTEC-2 studies have been published previously and are provided in the Supplemental Data [3, 4]. PORTEC-1 compared pelvic EBRT with no postoperative treatment in 715 women with intermediate risk, stage I EC [3]. PORTEC-2 compared postoperative EBRT with VBT in 427 women with high-intermediate risk stage I EC [4]. Cases were selected for this study based on the availability of tumor material for determination of molecular EC subgroups, and tissue microarrays (TMAs) for quantification of immune infiltrate. The study CONSORT diagram is provided as ([Supplemental Figure 1](#)).

Clinicopathological variables

Demographic variables and outcomes were obtained from the trial databases. Pathological analyses including criteria used to define LVSI and myometrial invasion were based on the results of central review, as reported previously [9, 11]. Four cases in which stage was altered to greater than stage IB following central review, but which were randomized, treated and reported in the PORTEC-2 study, were retained for analysis in this study. 16 cases identified as non-endometrioid endometrial cancer (NEEC) by central review were excluded from analysis, given the poor prognosis of this subtype. Full details of central pathology review are provided in the Supplemental Methods.

Determination of molecular subgroups and other molecular factors

Methods used to classify cases analogous to the TCGA subgroups have been reported previously [11, 24, 25]. Pathogenic *POLE* mutations were detected by sequencing *POLE* hotspot exons [24]. p53 status (abnormal staining/mutated – hereafter referred to as p53-

mutant vs. normal/wild-type) and MMR status (proficient vs. deficient) was determined by immunohistochemistry (IHC) in all cases, using previously reported methods [11]. *CTNNB1* mutation status was determined by next-generation sequencing (NGS) of exon 3. Cases with more than one classifying feature (e.g. *POLE* mutation and p53-mutant immunostaining pattern) were classified according to the dominant molecular feature on the basis of pathogenicity (e.g. MMRd cases with non-pathogenic *POLE* mutations were assigned to MMRd classifier) [26] (Supplemental Table 1). MMRp cases lacking pathogenic *POLE* mutation or p53 abnormality/mutation were assigned to the NSMP category. Methods used to determine additional molecular factors have been reported previously [11, 15] and are documented in the Supplemental Methods.

Machine learning based quantification and localization of immune infiltrate

TMAAs were produced as previously described [25]. Dual marker IHC for CD8 and CD103 was performed on TMA sections cut at 4µm using a modification of a previously reported protocol [27] (details in Supplemental Methods). Slides were scanned at high resolution (×200) and TMA cores segmented using the HALO™ digital image analysis software version v3.0.311.167 (Indica Labs, Corrales, NM, USA). Digital slide review and quality control was performed by an experienced pathologist (VHK). A deep neural network (DNN) algorithm (Simoyan and Zisserman VGG, HALO AI) was trained to localize and quantify tumor epithelial tissue and tumor-associated stroma regions, excluding non-tumor areas.

Classification accuracy was confirmed by review. Detection of cell nuclei and chromogens was performed by color deconvolution (Supplemental Methods), with thresholds set using internal controls. Marker-positive cells in stromal and epithelial regions in each core were classified, quantified at single cell resolution, and densities calculated as number of positive

cells/mm² [28]. Marker densities for each case were then calculated as the mean of the densities across all cores for that case.

Concordance of pathologist estimation of tumor CD8⁺ cell infiltrate with AI-based quantification

Analysis of concordance between AI-based quantification and expert pathologist review was calculated for total CD8⁺ cell counts in preference to densities, given the difficulty in accurately measuring surface area of tumor and stromal areas by visual inspection on the micrometer scale. A subset of 100 cases were reviewed by three pathologists who categorized them into four groups (0-5, 5-15, 15-50 and >50 cells per TMA core), broadly corresponding to the marker quartiles calculated across all cores. Further details are provided in Supplemental Methods.

Statistical analysis

Full details of statistical methods are provided in the Supplemental Methods. Hierarchical clustering was performed using Ward's minimum variance method [29, 30]. Biomarker analyses were performed in accordance with the REMARK guidelines [31] and are listed in Supplemental Table 2. For the analysis of immune infiltrate with clinical outcome, our primary endpoint was time to EC recurrence defined as the time from randomization to relapse, with censoring at last contact or death in case of no recurrence. The secondary endpoint was cancer-specific survival (CSS), defined as the time from randomization to EC death, with censoring at date of last contact or non-cancer death. Exploratory analyses are explicitly referred to as such. Survival curves were plotted using the Kaplan-Meier method. Time to event analyses were performed by pooled univariable and multivariable Cox proportional hazards models, stratified by trial. Covariables for inclusion in multivariable

models were pre-specified based on proven prognostic importance [11, 32], with number chosen to minimize risk of overfitting [31, 33]. Model validation was performed by bootstrap resampling [33]. Sensitivity analyses are detailed in Supplemental Methods. Concordance between categorization by pathologist review and AI-based quantification was calculated by weighted Cohen's kappa. Statistical analyses were performed using SPSS (Version 26) or R Version 3.6.1. (<http://www.r-project.org/>) (full list of packages provided in Supplemental Methods). All *P* values were two-sided. Statistical significance was accepted at $P < 0.05$.

Ethical approval

The PORTEC study protocols were approved by the Dutch Cancer Society and by the medical ethics committees at participating centers. Both studies were conducted in accordance with the principles of the Declaration of Helsinki. All patients provided signed informed consent to study participation.

RESULTS

Patient characteristics

The CONSORT diagram for this study is provided as Supplemental Figure 1, and characteristics of included patients are shown in Table 1. After exclusion of cases lacking tumor samples, failing QC or with non-endometrioid histology on central pathology review, 695 cases were informative for analysis (329 from PORTEC-1, and 366 from PORTEC-2), of which 691 (99.6%) were confirmed as stage I. Clinicopathological characteristics of these were similar to the original trial populations (<10% absolute difference in frequency of any variable), though comparison with excluded cases revealed modest, though statistically significant differences (Supplemental Table 3). Comparison of cases by trial reflected differences in the study inclusion criteria. The median follow up was 11.2 years (12.4 years in PORTEC-1 and 10.5 in PORTEC-2).

Immune cell density by clinicopathological factors and molecular EC subgroups

Increased density of tumor-infiltrating CD8⁺ cytotoxic T cells predicts favorable prognosis in multiple tumor types [34], and previous work has suggested that in EC it is the intraepithelial subset defined by co-expression of CD103 which confer prognostic value[35]. We therefore combined an image classification approach using deep neural networks with digital pathology methods quantifying both the density of single positive (CD8⁺ CD103⁻ and CD8⁻ CD103⁺) and double positive (CD8⁺ CD103⁺) cells, and their localization within the intraepithelial or intrastromal compartment (Figure 1A, Supplemental Methods, Supplemental Table 4). Analysis of the 695 cases revealed substantial variation in both the density of these cells and in that of total CD8⁺ cells and total CD103⁺ cells across tumors, with marked positive

skewness (3.56–5.88) (Supplemental Table 5, Supplemental Figure 2). As anticipated, the density of immune infiltrate was positively correlated between markers and compartments, though the strength of this relationship was variable (Spearman rho 0.13–0.76), being strongest for CD8⁺ single-positive and CD8⁺ CD103⁺ double-positive cells between the intraepithelial and intrastromal compartments (Figure 1B). Although the density of total CD8⁺ cells was lower in the intraepithelial than intrastromal compartment (mean 31.1 vs 41.9 cells/mm², $P=1.8 \times 10^{-5}$, Mann-Whitney test), the opposite was true of CD8⁺ CD103⁺ double positive cells (mean 20.4 vs 12.1 cells/mm², $P=1.8 \times 10^{-5}$), consistent with the known role of CD103 in intraepithelial localization.

Unsupervised hierarchical clustering of cases by immune marker density and compartment suggested three groups with high, intermediate and low immune cell infiltrate, with highly significant variation in the proportion of molecular EC subtypes between them ($P=8.3 \times 10^{-6}$) (Figure 1C, Supplemental Figure 3). The immune-high group contained more than three quarters of *POLE*-mutant cases (29 of 37; 78.4%), and a majority of the MMRd tumors (103 of 200; 51.5%), but proportionally fewer NSMP tumors (140 of 389; 36.0%). This group also contained the largest proportion of p53-mutant tumors (20 of 49; 40.8%); an unexpected finding given the poor prognosis of this group. In contrast, the immune-low group contained only 2 (5.4%) *POLE*-mutant tumors, but approximately one fifth of MMRd (40 of 200; 20%) and NSMP (90 of 389; 23.1%) cases. Interestingly, given recent data linking Wnt pathway activation with immune exclusion in tumors[36], mutation of *CTNNB1*, which encodes a key Wnt mediator was significantly enriched in the low (32 of 140; 22.9%) and intermediate (51 of 246; 20.7%) immune subgroups, compared to the high subgroup (41 of 284; 14.4%, $P=0.019$). Multiple linear regression revealed that *POLE* mutation and, to a lesser extent, MMRd were strongly predictive of immune infiltrate across all marker-compartment

combinations examined, while age, myometrial invasion, grade and *CTNNB1* mutation were less reliably associated, and LVSI, L1CAM overexpression, and p53 mutation showed no obvious association (Supplemental Table 6).

Prognostic value of immune cell infiltrate

We proceeded to examine the potential association of immune cell density with disease recurrence, mindful that this may vary by both cell type and compartment. We extended our previous analysis of CD8⁺ and CD103⁺ single positive cells and double positive cells to include total CD8⁺ cells (i.e. irrespective of CD103 status) and total CD103⁺ cells (i.e. irrespective of CD8 status) and to include analysis of the entire tumor region including both epithelial and stromal compartments. Univariable analysis of all 15 marker-compartment combinations in the pooled trial population with internal bootstrap validation (n=1000) revealed that CD8⁺ CD103⁻ single positive cells in either the epithelial compartment or combined epithelial and stromal compartments, and intraepithelial total CD8⁺ density were the strongest predictors of tumor recurrence (Supplemental Table 7). Of these, intraepithelial CD8⁺ density was selected for further analysis on the basis of its effect size (HR=0.88 per two-fold increase; 95% CI=0.82–0.95; *P*=0.001), Akaike information criterion (AIC) and clinical applicability of a single immunostain. Interestingly, and in contrast to other malignancies [37], intrastromal immune infiltrate had no discernible prognostic value.

Integration of molecular EC classification and image-based immunoprofiling

We next examined whether quantification of intraepithelial total CD8⁺ cell density could further enhance risk stratification beyond the improvement the molecular EC classification

(including L1CAM positivity) provides over clinicopathological variables alone [11]. We first confirmed that addition of these molecular factors improved prediction of disease recurrence and goodness of model fit compared to a ‘pathological’ multivariable model containing clinicopathological variables only (LR test $P=1.2 \times 10^{-6}$) (Supplemental Table 8). Addition of intraepithelial CD8⁺ cell density to this ‘molecular’ model further improved model fit (LR test $P=3.0 \times 10^{-3}$), and confirmed the independent prognostic value of this marker, with multivariable-adjusted HR of 0.88 for each two-fold increase (95% CI=0.81–0.95; $P=1.8 \times 10^{-3}$) (Table 2, Figure 3A). This association was essentially unchanged after exclusion of the four (0.6%) cases reclassified as stage II/IIIA by central review, by the inclusion of *CTNNB1* mutation in the multivariable model, or by exclusion of grade 3 and *POLE*-mutant cases (Supplemental Tables 9-11). The prognostic value of intraepithelial CD8⁺ density equated to a multivariable-adjusted HR of 0.70 (95% CI=0.57–0.88) for comparison of cases with a density at the 75th percentile with vs those at the 25th percentile, and a HR of 0.50 (95% CI=0.33–0.78) for cases at the 90th percentile vs those at the 10th percentile. Corresponding point estimates of the likelihood of being recurrence-free at 3 years were 88.1% (95% CI=84.0–92.3%), 89.9% (86.7–93.3%), 92.6% (89.7–95.6%) and 93.5% (90.6–96.5%), for cases at the 10th, 25th, 75th, and 90th percentile of CD8⁺ cell density respectively. The prognostic value of intraepithelial CD8⁺ cell density was also observed when dichotomized at the sample median (multivariable-adjusted HR for high vs low=0.50; 95% CI=0.31-0.81; $P=5.1 \times 10^{-3}$) (Figure 3B). Sensitivity analysis showed no evidence that these results were due to differential effect of radiotherapy by CD8⁺ cell density (Supplemental Table 12).

Comparison of multivariable models with internal bootstrap validation demonstrated that the molecular-immune model had the highest concordance, confirming the independent

prognostic value of both the molecular EC classification and immune infiltrate (Table 2, Figure 2C). Further analysis of this molecular-immune model revealed that intraepithelial CD8⁺ cell density was a more important predictor of recurrence than myometrial invasion, tumor grade, or L1CAM positivity (Figure 2D). Intraepithelial CD8⁺ cell infiltrate was also predictive of endometrial cancer-specific survival in multivariable analysis both as a continuous variable (HR=0.89; 95% CI=0.81–0.98; *P*=0.015) and when dichotomized (HR=0.47, 95% CI=0.26–0.87, *P*=0.016) (Table 2, Figure S3).

Prognostic value of intraepithelial CD8⁺ infiltrate within TCGA subgroups and by *CTNNB1* mutation

The unexpectedly high prevalence of p53-mutant cases in the immune high cluster, and the lack of association of MMRd with reduced recurrence despite typically prominent T cell infiltrate motivated us to explore the prognostic value of intraepithelial CD8⁺ infiltrate within EC molecular subgroups. CD8⁺ cell density was a statistically significant predictor of recurrence in the p53-mutant subgroup (*n*=47) in univariable analysis (HR=0.84; 95% CI=0.73–0.97; *P*=0.017), and after adjusting for grade and LVSI in multivariable analysis (HR=0.84; 95% CI=0.73–0.98; *P*=0.024) (Table 3, Figure 3A). A similar, albeit weaker association was evident in the NSMP subgroup (*n*=374), although this fell just outside the margin of statistical significance in multivariable analysis (univariable HR=0.88; 95% CI=0.77–1.00; *P*=0.048, adjusted HR=0.88; 95% CI=0.77–1.00; *P*=0.055 respectively) (Table 3, Figure 3A). In contrast, among MMRd tumors (*n*=186), CD8⁺ infiltrate had no detectable prognostic value (univariable HR= 0.98; 95% CI=0.84–1.16; *P*=0.85, adjusted HR=0.95; 95% CI=0.79–1.13; *P*=0.56). The small number of cases and events in the POLE-mutant subgroup precluded similar analysis (Table 3, Figure 3B). Further exploratory analysis by

CTNNB1 mutation revealed CD8⁺ density held similar prognostic value among wild-type cases as the total study population (HR=0.86, 95% CI=0.79-0.94, $P=1.1\times10^{-3}$); the modest size of the *CTNNB1*-mutant subgroup precluded firm conclusions (Supplemental Table 13).

Concordance between AI-based quantification and pathologist-based estimation of tumor CD8⁺ cell infiltration

While CD8 immunohistochemistry is a standard assay in pathology laboratories, AI-based image analysis and quantification is, at present, limited to the research setting. We therefore sought to determine whether a simple pathologist estimation of intraepithelial CD8⁺ cell infiltrate could serve as a surrogate for AI-based quantification for future study. Taking CD8⁺ cell counts (i.e. the number of intraepithelial cells within each TMA core) as a measure readily amenable to pathologist estimation (in contrast to density, which requires determination of the area under analysis), we defined four groups, broadly corresponding to the quartiles determined by AI-based analysis (0-5, 5-15, 15-50 and >50 cells per TMA core). A subset of 100 cases were then independently reviewed by three pathologists, and assigned to these groups, each blinded to the results of the others and the AI-based analysis. Analysis of the results revealed substantial to strong concordance between individual pathologist categorisation (weighted Cohen's kappa 0.70–0.89, all $P<0.001$), and moderate to substantial concordance between individual and consensus pathological categorisation and the AI-based algorithm (kappa 0.4–0.65) (Supplemental Table 14). On review, this difference was determined to be due to systematic underscoring by the pathologists compared to the AI-based method; i.e. the AI-based method demonstrated greater sensitivity. While pathologist evaluation of the entire study population was beyond the scope of this study, these could serve as pragmatic cut points for future validation.

DISCUSSION

In this study, we show that machine learning, image-based quantification of intraepithelial CD8⁺ cells refines prognostication in early-stage endometrioid EC beyond clinicopathological and molecular factors. Our multimodal approach holds promise to improve risk stratification, and thus reduce over- and undertreatment in this common cancer.

While previous studies have shown the strong prognostic value of the molecular EC classification [11, 16], and of the density of tumor-infiltrating lymphocytes in EC [19, 20], to our knowledge our study is one of only two to examine the combination of these factors. The other study, by Talhouk and colleagues [23], reported no prognostic effect of intraepithelial CD3⁺ CD8⁺ cells in multivariable analysis which included the molecular EC classification, although a tendency to improved survival was observed. The discordance with our results is currently unexplained, but may relate to the methodology used for immune cell localization, statistical analysis, or the smaller, non-trial population of mixed histotypes and stages used in their study.

Interestingly, and also in contrast to Talhouk and colleagues' results [23], our data suggest that image-based assessment of immune infiltrate has particular value in specific molecular subgroups. The first are low mutation burden NSMP tumors, which constitute the majority of early-stage endometrioid EC and which currently lack reliable molecular stratifiers. The second are low mutation burden p53-mutant cases, which have poor prognosis and are increasing in prevalence [38]. In contrast, we found no evidence that intraepithelial CD8⁺ cell density was prognostic in high mutation burden MMRd tumours, despite their overall enhanced T-cell infiltrate compared to NSMP and p53-mutant subgroups. The reasons for

this discrepancy are unclear. While it is tempting to speculate on the antigenicity of somatic copy number alterations (SCNAs) in p53-mutant tumors, or attenuation of T cell-mediated cytotoxicity by immune checkpoint upregulation in MMRd cases, these exploratory analyses require validation before conclusions can be drawn and mechanisms investigated.

Strengths of our study include its large size, homogeneous clinical trial cohorts with meticulous follow-up data [3, 4], central pathological review and comprehensive annotation of pathological and molecular risk factors [11]. Another strength is our combination of machine-learning based image segmentation methods with digital pathology to enable automated analysis of the type, density, and localization of T cells – the importance of which is illustrated by the differing prognostic value of intraepithelial and intrastromal infiltrates. Our approach is readily implementable with a single immunostain in routine clinical use, and thus represents an advance on previous methods, which rely on manual cell counting or multispectral analysis, neither of which are deliverable at the scale required for clinical practice.

Our study has limitations. For logistical reasons, we used TMAs rather than whole tissue sections. Given the potential for intratumoral heterogeneity in both molecular alterations (e.g. subclonal MMR loss) and immune infiltrate in tumors, it will be important to determine whether our results could be improved by analysis of a larger tissue area, ideally including the tumor invasive margin. Similar considerations limited our focus mainly to cytotoxic T cells; it will also be of interest to see whether analysis of additional immune cell types improves prognostication, notwithstanding the issues of scalability noted above. Finally, while chemotherapy was not used in our study population, it is theoretically possible that our

results could reflect enhanced radiosensitivity in CD8⁺ high tumors, although sensitivity analysis did not support this.

While our results illustrate the potential of digital pathology to improve clinical care, the infrastructure required to implement this is currently limited to specialist centers. We therefore evaluated whether a simple categorization of cases by pathologists according to CD8⁺ cell infiltrate could serve as a surrogate for the AI-based quantification. The good concordance between methods suggests that, with validation, the cut points we define could be used in clinical practice in the near-term. However, the pragmatic, semi-quantitative pathologist-based assessment was generally less sensitive than the AI-based classifier. This highlights that cut-offs for immune cell assessment are method-dependent and underscores the advantage of deriving continuous scores using unbiased, automated methods for image analysis at single cell resolution.

To conclude, we show that image-based quantification of intraepithelial CD8⁺ cells improves the strength of the molecular EC classification and refines prognostication in early-stage endometrioid EC. Future work will seek to validate this finding, particularly in higher-risk cohorts including non-endometrioid histotypes such as the PORTEC-3 study [6], and to validate this novel multimodal approach for clinical implementation.

AUTHOR CONTRIBUTIONS

Study design: NH, MB, TB, RN, VK, DNC

Data collection: all authors

Data analysis: NH, MB, AL, TB, RN, VK, DNC

Manuscript writing: NH, DNC

Manuscript contribution and approval: all authors

ACKNOWLEDGEMENTS

We are grateful to the participants in the PORTEC-1 and PORTEC-2 trial, and to the investigators and pathologists who recruited patients and collected samples.

REFERENCES

- 1 Ferlay J, Steliarova-Foucher E, Lortet-Tieulent J, Rosso S, Coebergh JW, Comber H et al. Cancer incidence and mortality patterns in Europe: estimates for 40 countries in 2012. *Eur J Cancer* **49**, 1374-1403 (2013).
- 2 Siegel RL, Miller KD & Jemal A Cancer statistics, 2019. *CA: A Cancer Journal for Clinicians* **69**, 7-34 (2019).
- 3 Creutzberg CL, van Putten WL, Koper PC, Lybeert ML, Jobsen JJ, Warlam-Rodenhuis CC et al. Surgery and postoperative radiotherapy versus surgery alone for patients with stage-1 endometrial carcinoma: multicentre randomised trial. PORTEC Study Group. Post Operative Radiation Therapy in Endometrial Carcinoma. *Lancet* **355**, 1404-1411 (2000).
- 4 Nout RA, Smit VT, Putter H, Jurgensliemk-Schulz IM, Jobsen JJ, Lutgens LC, et al. Vaginal brachytherapy versus pelvic external beam radiotherapy for patients with endometrial cancer of high-intermediate risk (PORTEC-2): an open-label, non-inferiority, randomised trial. *Lancet* **375**, 816-823 (2010).
- 5 Wortman BG, Creutzberg CL, Putter H, Jurgensliemk-Schulz IM, Jobsen JJ, Lutgens L. et al. Ten-year results of the PORTEC-2 trial for high-intermediate risk endometrial carcinoma: improving patient selection for adjuvant therapy. *Br J Cancer* **119**, 1067-1074 (2018).
- 6 de Boer SM, Powell ME, Mileschkin L, Katsaros D, Bessette P, Haie-Meder C. et al. Adjuvant chemoradiotherapy versus radiotherapy alone for women with high-risk endometrial cancer (PORTEC-3): final results of an international, open-label, multicentre, randomised, phase 3 trial. *The Lancet. Oncology* **19**, 295-309 (2018).
- 7 Creutzberg CL, Lu KH & Fleming GF. Uterine Cancer: Adjuvant Therapy and Management of Metastatic Disease. *J Clin Oncol* **37**, 2490-2500 (2019).
- 8 Colombo N, Creutzberg C, Amant F, Bosse T, Gonzalez-Martin A, Ledermann J. et al. ESMO-ESGO-ESTRO Consensus Conference on Endometrial Cancer: diagnosis, treatment and follow-up dagger. *Ann Oncol* **27**, 16-41 (2016).
- 9 Bosse T, Peters EE, Creutzberg CL, Jurgensliemk-Schulz IM, Jobsen JJ, Mens JW, et al. Substantial lympho-vascular space invasion (LVSI) is a significant risk factor for recurrence in endometrial cancer--A pooled analysis of PORTEC 1 and 2 trials. *Eur J Cancer* **51**, 1742-1750 (2015).

- 10 Kurnit KC, Kim GN, Fellman BM, Urbauer DL, Mills GB, Zhang W. et al. CTNNB1 (beta-catenin) mutation identifies low grade, early stage endometrial cancer patients at increased risk of recurrence. *Modern Pathology* **30**, 1032-1041 (2017).
- 11 Stello E, Nout RA, Osse EM, Jurgenliemk-Schulz IJ, Jobsen JJ, Lutgens LC, et al. Improved Risk Assessment by Integrating Molecular and Clinicopathological Factors in Early-stage Endometrial Cancer- Combined Analysis of the PORTEC Cohorts. *Clin Cancer Res* **22**, 4215-4224, (2016).
- 12 Zeimet AG, Reimer D, Huszar M, Winterhoff B, Puistola U, Azim SA et al.. L1CAM in early-stage type I endometrial cancer: results of a large multicenter evaluation. *J Natl Cancer Inst* **105**, 1142-1150, (2013).
- 13 Liu Y, Patel L, Mills GB, Lu KH, Sood AK, Ding L, et al. Clinical significance of CTNNB1 mutation and Wnt pathway activation in endometrioid endometrial carcinoma. *J Natl Cancer Inst* **106** (2014).
- 14 The Cancer Genome Atlas. Integrated genomic characterization of endometrial carcinoma. *Nature* **497**, 67-73 (2013).
- 15 Bosse T, Nout RA, Stelloo E, Dreef E, Nijman HW, Jurgenliemk-Schulz IM, et al. L1 cell adhesion molecule is a strong predictor for distant recurrence and overall survival in early stage endometrial cancer: pooled PORTEC trial results. *Eur J Cancer* **50**, 2602-2610, (2014).
- 16 Kommoss S, McConechy MK, Kommoss F, Leung S, Bunz A, Magrill J, et al. Final validation of the ProMisE molecular classifier for endometrial carcinoma in a large population-based case series. *Ann Oncol* **29**, 1180-1188 (2018).
- 17 Costigan DC, Dong F, Nucci MR & Howitt BE. Clinicopathologic and Immunohistochemical Correlates of CTNNB1 Mutated Endometrial Endometrioid Carcinoma. *Int J Gynecol Pathol* **39**, 119-127 (2020).
- 18 Wortman BG, Bosse T, Nout RA, Lutgens L, van der Steen-Banasik EM, Westerveld H et al. Molecular-integrated risk profile to determine adjuvant radiotherapy in endometrial cancer: Evaluation of the pilot phase of the PORTEC-4a trial. *Gynecol Oncol* **151**, 69-75, (2018).
- 19 Kondratiev S, Sabo E, Yakirevich E, Lavie O & Resnick MB. Intratumoral CD8+ T lymphocytes as a prognostic factor of survival in endometrial carcinoma. *Clin Cancer Res* **10**, 4450-4456, (2004).
- 20 de Jong RA, Leffers N, Boezen HM, ten Hoor KA, van der Zee AG, Hollema H et al. Presence of tumor-infiltrating lymphocytes is an independent prognostic factor in type I and II endometrial cancer. *Gynecol Oncol* **114**, 105-110 (2009).

- 21 van Gool IC, Eggink FA, Freeman-Mills L, Stelloo E, Marchi E, de Bruyn M et al. POLE Proofreading Mutations Elicit an Antitumor Immune Response in Endometrial Cancer. *Clin Cancer Res* **21**, 3347-3355, (2015).
- 22 Eggink FA, Van Gool IC, Leary A, Pollock PM, Crosbie EJ, Mileskin L et al. Immunological profiling of molecularly classified high-risk endometrial cancers identifies POLE-mutant and microsatellite unstable carcinomas as candidates for checkpoint inhibition. *Oncoimmunology* **6**, e1264565, (2017).
- 23 Talhouk A, Derocher H, Schmidt P, Leung S, Milne K, Gilks CB, et al. Molecular Subtype Not Immune Response Drives Outcomes in Endometrial Carcinoma. *Clin Cancer Res* **25**, 2537-2548, (2019).
- 24 Church DN, Stelloo E, Nout RA, Valtcheva N, Depreeuw J, ter Haar N et al. Prognostic significance of POLE proofreading mutations in endometrial cancer. *J Natl Cancer Inst* **107**, 402, doi:10.1093/jnci/dju402 (2015).
- 25 Stelloo E, Bosse T, Nout RA, MacKay HJ, Church DN, Nijman HW et al. Refining prognosis and identifying targetable pathways for high-risk endometrial cancer; a TransPORTEC initiative. *Modern pathology* **28**, 836-844, (2015).
- 26 León-Castillo A, Gilvazquez E, Nout R, Smit VT, McAlpine JN, McConechy M et al. Clinicopathological and molecular characterisation of 'multiple-classifier' endometrial carcinomas. *J Pathol* **250**, 312-322, (2020).
- 27 Workel HH, Lubbers JM, Arnold R, Prins TM, van der Vlies P, de Lange K et al. Transcriptionally Distinct CXCL13(+)CD103(+)CD8(+) T-cell Population Is Associated with B-cell Recruitment and Neoantigen Load in Human Cancer. *Cancer Immunology Research* **7**, 784-796, (2019).
- 28 Koelzer VH, Sirinukunwattana K, Rittscher J, & Mertz KD. Precision immunoprofiling by image analysis and artificial intelligence. *Virchows Archiv* **474**, 511-522, (2019).
- 29 Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*. (Springer Verlag, 2009).
- 30 Murtagh F & Legendre P. Ward's Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward's Criterion? *J Classification* **31**, 274–295 (2014)
- 31 McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM. Reporting Recommendations for Tumor Marker Prognostic Studies (REMARK). *J Nat Cancer Inst* **97**, 1180-1184 (2005).
- 32 Creutzberg, CL, van Stiphout RG, Nout RA, Lutgens LC, Jurgenliemk-Schulz IM, Jobsen JJ et al. Nomograms for prediction of outcome with or without adjuvant radiation therapy for patients with endometrial cancer: a pooled analysis of PORTEC-1 and PORTEC-2 trials. *Int J Rad Oncol, Biol, Phy* **91**, 530-539, (2015).

- 33 Harrell FE *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. (Springer, 2001).
- 34 Fridman WH, Pages F, Sautes-Fridman C & Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer* **12**, 298-306, (2012).
- 35 Workel HH, Komdeur FL, Wouters MC, Plat A, Klip HG, Eggink FA et al. CD103 defines intraepithelial CD8+ PD1+ tumour-infiltrating lymphocytes of prognostic significance in endometrial adenocarcinoma. *Eur J Cancer* **60**, 1-11, (2016).
- 36 Luke JJ, Bao R, Sweis RF, Spranger S & Gajewski TF WNT/beta-catenin Pathway Activation Correlates with Immune Exclusion across Human Cancers. *Clin Cancer Res* **25**, 3074-3083, (2019).
- 37 Salgado R, Denkert C, Demaria S, Sirtaine N, Klauschen F, Pruneri G et al. The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014. *Ann Oncol* **26**, 259-271, (2015).
- 38 Clarke MA, Devesa SS, Harvey SV & Wentzensen N. Hysterectomy-Corrected Uterine Corpus Cancer Incidence Trends and Differences in Relative Survival Reveal Racial Disparities and Rising Rates of Nonendometrioid Cancers. *J Clin Oncol* **37**, 1895-1908, (2019).

TABLES

Table 1. Characteristics of study participants

	PORTEC-1	PORTEC-2	Total	
Characteristics	n=329	n=366	n=695	P value
Age: median (IQR)	67 (13)	66 (12)	70 (10)	<0.0001
Stage (2009 classification)				
IA	126 (38.3%)	59 (16.1%)	185 (26.6%)	<0.0001
IB	203 (61.7%)	303 (82.8%)	506 (72.8%)	
>IB*	0 (0.0%)	4 (1.0%)	4 (0.6%)	
Grade				
1	225 (68.4%)	292 (79.8%)	517 (74.4%)	0.014†
2	51 (15.5%)	39 (10.7%)	90 (12.9%)	
3	53 (16.1%)	35 (9.6%)	88 (12.7%)	
Myometrial invasion				
≤50%	126 (38.3%)	58 (15.8%)	184 (26.5%)	<0.0001
>50%	203 (61.7%)	308 (84.2%)	511 (73.5%)	
LVSI (none-mild vs severe)				
None	278 (84.5%)	282 (77.0%)	560 (80.6%)	0.86‡
Mild	22 (6.7%)	47 (12.8%)	69 (9.9%)	
Severe	14 (4.3%)	17 (4.6%)	31 (4.5%)	
Unknown	15 (4.6%)	20 (5.5%)	35 (5.0%)	
LICAM				
None or ≤10% positive cells	313 (95.1%)	342 (93.4%)	655 (94.2%)	0.41
>10% positive cells	16 (4.9%)	23 (6.3%)	39 (5.6%)	
Unknown	0 (0.0%)	1 (0.3%)	1 (0.1%)	
Molecular group				
NSMP	177 (55.3%)	212 (59.7%)	389 (57.6%)	0.79
POLE	20 (6.3%)	17 (4.8%)	37 (5.5%)	
MMRd	98 (30.6%)	102 (28.7%)	200 (29.6%)	
p53-mutant	25 (7.8%)	24 (6.8%)	49 (7.3%)	
Unknown	9 (2.7%)	11 (3.0%)	20 (2.9%)	
Received adjuvant treatment§				
None	172 (52.3%)	3 (0.8%)	175 (25.5%)	<0.0001
Vaginal brachytherapy	0 (0.0%)	184 (50.3%)	336 (48.3%)	
Pelvic EBRT	157 (47.7%)	179 (48.9%)	184 (26.5%)	

Comparison of cases included in this biomarker study with those excluded and the total trial populations are provided as Supplemental Table 2. *Includes four cases for which staging was revised following central pathological review: two cases as Stage II, and two cases as stage IIIA (see Supplemental Methods for details). † comparison of proportion of grade 3 vs grade 1-2. ‡ comparison of severe LVSI vs none or mild. § adjuvant chemotherapy given in either study. IQR – interquartile range; LVSI – lymphovascular space invasion; LICAM – L1 cell adhesion molecule; EC – endometrial cancer; NSMP – no specific molecular profile; POLE – pathogenic POLE exonuclease domain mutant; MMRd – DNA mismatch repair deficient; p53-abnormal – mutant p53 staining pattern on immunohistochemistry or *TP53* mutation; EBRT – external beam radiotherapy.

Table 2. Univariable and multivariable analyses of time to endometrial cancer recurrence and endometrial cancer specific survival in pooled PORTEC-1 and PORTEC-2 trial population

	Univariable analysis		Multivariable analysis	
	HR (95% CI)	<i>P</i>	HR (95% CI)	<i>P</i>
Endometrial cancer recurrence (640 cases, 80 events)				
Myometrial invasion				
≤50%	1.0 (ref)	–	1.0 (ref)	–
>50%	1.46 (0.88–2.43)	0.15	2.39 (1.32–4.30)	3.9 x 10 ⁻³
Grade				
1-2	1.0 (ref)	–	1.0 (ref)	–
3	2.49 (1.53–4.03)	0.00022	2.21 (1.25–3.91)	6.4 x 10 ⁻³
LVSI				
Absent/mild	1.0 (ref)	–	1.0 (ref)	–
Severe	3.25 (1.67–6.30)	0.00051	4.19 (2.09–8.38)	5.1 x 10 ⁻⁵
L1CAM				
≤10% staining	1.0 (ref)	–	1.0 (ref)	–
>10% staining	4.70 (2.69–8.21)	5.5 x 10 ⁻⁸	2.26 (1.12–4.53)	0.022
Molecular EC group				
NSMP	1.0 (ref)	–	1.0 (ref)	–
POLE	0.80 (0.25–2.58)	0.71	1.41 (0.41–4.75)	0.59
MMRd	1.43 (0.88–2.33)	0.15	1.30 (0.75–2.23)	0.35
p53-mutant	4.84 (2.76–8.48)	3.5 x 10 ⁻⁸	4.98 (2.51–9.88)	4.2 x 10 ⁻⁶
Intraepithelial CD8 ⁺ cell density (continuous, per doubling)	0.89 (0.82–0.96)	0.0036	0.88 (0.81–0.95)	1.9 x 10 ⁻³
Endometrial cancer-specific survival (640 cases, 53 events)				
Myometrial invasion				
≤50%	1.0 (ref)	–	1.0 (ref)	–
>50%	1.20 (0.66–2.19)	0.55	2.22 (1.10–4.48)	0.026
Grade				
1-2	1.0 (ref)	–	1.0 (ref)	–
3	3.50 (2.03–6.04)	6.6 x 10 ⁻⁶	2.58 (0.33–4.99)	5.1 x 10 ⁻³
LVSI				
Absent/mild	1.0 (ref)	–	1.0 (ref)	–
Severe	3.94 (1.86–8.37)	0.00035	5.33 (2.37–11.98)	5.2 x 10 ⁻⁵
L1CAM				
≤10% staining	1.0 (ref)	–	1.0 (ref)	–
>10% staining	5.19 (2.75–9.79)	3.6 x 10 ⁻⁷	2.26 (1.03–4.96)	0.043
Molecular EC group				
NSMP	1.0 (ref)	–	1.0 (ref)	–
POLE	1.01 (0.24–4.32)	0.99	1.70 (0.38–7.62)	0.49
MMRd	2.12 (1.16–3.86)	0.014	1.68 (0.84–3.34)	0.14
p53-mutant	7.35 (3.78–14.3)	4.0 x 10 ⁻⁹	6.79 (3.3–15.22)	3.2 x 10 ⁻⁶
Intraepithelial CD8 ⁺ cell density (continuous, per doubling)	0.88 (0.80–0.97)	0.0077	0.89 (0.81–0.98)	0.015

Cox models use all informative cases and exclude those with missing data in case of multivariable models (maximum 5.0% missing data for any variable). Multivariable models included prespecified covariables of known prognostic value (see Methods) and were not subject to variable selection. Results from corresponding Cox models for endometrial cancer recurrence and endometrial cancer-specific survival before and after addition of CD8⁺ cell density are provided in Supplemental Table S5. Results from analysis after exclusion of the four (0.6%) cases classified as > stage I on central review are shown in Supplemental Table S6. The addition of intraepithelial CD8⁺ cell density to the molecular model (containing clinicopathological variables, the molecular EC classifier and L1CAM) for endometrial cancer recurrence was associated with an improvement in model fit evidenced by: (i) reduction in Akaike Information Criterion [AIC] (molecular model vs. integrated molecular-immune model = 842.2 vs. 835.4); (ii) increase in model concordance (C index 0.697 vs. 0.726); and (iii) Likelihood ratio test for comparison of nested models: $P=3.0 \times 10^{-3}$). Similar, albeit less strong, results were obtained from comparison of models for endometrial cancer-specific survival (molecular model vs. molecular-immune model): (i) AIC=544.9 vs. 541.4; (ii) C index 0.792 vs. 0.802; (iii) likelihood ratio test for comparison of models: $P=0.019$. HR – hazard ratio; 95% CI – 95% confidence interval; LVSI – lymphovascular space invasion; L1CAM – L1 cell adhesion molecule; EC – endometrial cancer; NSMP – no specific molecular profile; POLE – pathogenic POLE exonuclease domain mutant; MMRd – DNA mismatch repair deficient; p53-mutant – mutant p53 staining pattern on immunohistochemistry or *TP53* mutation.

Table 3. Exploratory univariable and multivariable analyses of time to endometrial cancer recurrence according to molecular endometrial cancer subgroup

	Univariable analysis		Multivariable analysis	
	HR (95% CI)	<i>P</i>	HR (95% CI)	<i>P</i>
NSMP (374 cases, 35 events)				
Grade				
1-2	1.0 (ref)	–	1.0 (ref)	–
3	4.55 (1.95–10.65)	4.7 x 10 ⁻⁴	5.61 (2.23 – 14.13)	2.5 x 10 ⁻⁴
LVSI				
Absent/mild	1.0 (ref)	–	1.0 (ref)	–
Severe	3.14 (0.95–10.45)	0.062	3.09 (0.92 – 10.41)	0.068
CD8 ⁺ cell density (continuous)	0.88 (0.77–0.999)	0.048	0.88 (0.77 – 1.00)	0.055
POLE (34 cases, 3 events)				
	Not done	–	Not done	–
MMRd (186 cases, 24 events)				
Grade				
1-2	1.0 (ref)	–	1.0 (ref)	–
3	1.43 (0.60–3.37)	0.41	1.16 (0.45 – 3.01)	0.75
LVSI				
Absent/mild	1.0 (ref)	–	1.0 (ref)	–
Severe	4.97 (2.03–12.13)	4.4 x 10 ⁻⁴	4.91 (1.96 – 12.3)	6.8 x 10 ⁻⁴
CD8 ⁺ cell density (continuous)	0.98 (0.84–1.16)	0.85	0.95 (0.79 – 1.13)	0.56
p53-mutant (47 cases, 18 events)				
Grade				
1-2	1.0 (ref)	–	1.0 (ref)	–
3	1.39 (0.54–3.56)	0.49	1.12 (0.41 – 3.04)	0.82
LVSI*	–	–	–	–
CD8 ⁺ cell density (continuous)	0.84 (0.73–0.97)	0.017	0.84 (0.73 – 0.98)	0.024

Univariable and multivariable hazard ratios are derived from complete case analyses and exclude cases with missing data (maximum 7.0% for any single variable). Analysis of POLE subgroup was not performed due to the very small number of events within this subset.. Covariables of tumor grade and LVSI used in multivariable analyses were prespecified based upon prognostic value in pooled study population; variable selection was not performed. *Failure of model convergence precluded analysis of LVSI within the p53-abnormal subgroup. HR – hazard ratio; 95% CI – 95% confidence interval; LVSI – lymphovascular space invasion; L1CAM – L1 cell adhesion molecule; NSMP – no specific molecular profile; POLE – pathogenic POLE exonuclease domain mutant; MMRd – DNA mismatch repair deficient; p53-mutant – mutant p53 staining pattern on immunohistochemistry or *TP53* mutation.

FIGURE LEGENDS

Figure 1. CD8⁺ and CD103⁺ cell density by location and relationship with molecular factors

(A) Following dual marker immunohistochemistry (IHC), an image based, machine learning algorithm was developed to quantify the density of immune cells expressing the cytotoxic T cell marker CD8 and/or the intraepithelial T cell marker CD103 within the intraepithelial and intrastromal compartments. In addition to CD8⁺ CD103⁻ cells (red arrows), CD8⁻ CD103⁺ cells (yellow arrows) and CD8⁺ CD103⁺ cells (orange arrowheads), we also analyzed the total number of CD8⁺ cells (irrespective of CD103 status) and total number of CD103⁺ cells (irrespective of CD8 status) in these regions separately and in combination (intratumoural density), to give a total of 15 marker-compartment combinations. (B) Matrix showing Spearman correlation between cell populations by cell surface markers and localization. Densities of total CD8⁺ and total CD103⁺ cells in all compartments, and all intratumoural marker densities were excluded as these are determined by the densities of subpopulations and within subregions. (C) Unsupervised hierarchical clustering of immune cell densities following mean centering and scaling, using identical marker-compartment combinations as (B). Cluster number (i.e. dendrogram cut height) was selected based on the results of the gap statistic. Bars to the right of the heatmap show density of intratumoural and intraepithelial CD8⁺ cell density, TCGA molecular subgroup and CTNNB1 mutation status. Boxplots below heatmap indicate cell density by TCGA subgroup for each marker-compartment combination (shown in full in Supplemental Figure 3). Lower and upper limits of box indicate 25th and 75th percentiles; whiskers extend to 1.5x interquartile range below and above these values respectively; horizontal line within box indicates median. DNN – deep neural network; NSMP – no specific molecular profile; POLE – pathogenic POLE

exonuclease domain mutant; MMRd – DNA mismatch repair deficient; p53-abnormal – mutant p53 staining pattern on immunohistochemistry or *TP53* mutation.

Figure 2. Prognostic value of intraepithelial total CD8⁺ infiltrate in pooled study population

(A) Upper: plot showing hazard ratio for endometrial cancer recurrence according to (log2 transformed) density of intraepithelial total CD8⁺ cells, after adjusting for covariables. Note that each unit increase corresponds to a doubling in marker density. Lower: corresponding kernel density plot showing proportion of cases according to density of immune infiltrate. Dashed vertical lines correspond to 5th, 25th, 50th, 75th and 95th percentiles. (B) Kaplan Meier curves showing time to endometrial cancer recurrence by intraepithelial total CD8⁺ density divided at the sample median. (C) Boxplots showing concordance (C index) of base model (pathological factors only), molecular model (pathological factors plus molecular EC classifier and L1CAM expression) and integrated molecular-immune model (molecular model plus intraepithelial total CD8⁺ density for time to endometrial cancer recurrence. Box and whisker (Tukey) plots use results of 1000 bootstrap resamples from study population; lower and upper limits of box indicate 25th and 75th percentiles; whiskers extend to 1.5x interquartile range below and above these values respectively. Thick vertical colored lines within box indicate median value from bootstrap resamples; dashed vertical line indicates C index from original biomarker population. (D) Pie charts showing relative importance of variables within these three multivariable models based on the proportion of the χ^2 statistic. *P* value in (B) was determined by the log-rank test; shaded area indicates 95% confidence interval. NSMP – no specific molecular profile; POLE – pathogenic POLE exonuclease

domain mutant; MMRd – DNA mismatch repair deficient; p53-mutant – mutant p53 staining pattern on immunohistochemistry or *TP53* mutation.

Figure 3. Prognostic value of intraepithelial CD8⁺ cell infiltrate within molecular endometrial cancer subgroups

Kaplan-Meier curves showing probability of endometrial cancer recurrence according to intraepithelial total CD8⁺ cell density (dichotomized at median of study population) for (A) NSMP and p53-mutant tumors, characterized by low mutation burden, and (B) MMRd and POLE tumors, characterized by high mutation burden (POLE subgroup was not subdivided as only three cases had CD8⁺ cell densities below the study median). *P* values were calculated by the log-rank test; shaded area indicates 95% confidence interval. CD8⁺ – intraepithelial total CD8⁺ cell density; NSMP – no specific molecular profile; POLE – pathogenic POLE exonuclease domain mutant; MMRd – DNA mismatch repair deficient; p53-mutant – mutant p53 staining pattern on immunohistochemistry or *TP53* mutation.