Performance of Next-Generation Sequencing for the Detection of Microsatellite Instability in Colorectal Cancer With Deficient DNA Mismatch Repair

Toky Ratovomanana, Romain Cohen, Magali Svrcek, Florence Renaud, Pascale Cervera, Aurélie Siret, Quentin Letourneur, Olivier Buhard, Pierre Bourgoin, Erell Guillerm, et al.

Version of Record: https://www.sciencedirect.com/science/article/pii/S0016508521029772
Manuscript_1632f3740930e940781e98ba824408f1

Performance of Next Generation Sequencing for the Detection of Microsatellite Instability in Colorectal Cancer with Deficient DNA Mismatch Repair

Short Title: Assessment of NGS for MSI diagnosis in CRC

Toky Ratovomanana[1], Romain Cohen[1,3], Magali Svrcek[1,4], Florence Renaud[1,9], Pascale Cervera[1,2], Aurélie Siret[1], Quentin Letourneur[1], Olivier Buhard[1], Pierre Bourgoin[1], Erell Guillerm[1,2], Coralie Dorard[1], Remy Nicolle[5], Mira Ayadi[5], Mehdi Touat[1,6], Franck Bielle[6,8], Marc Sanson[6,8], Philippe Le Rouzic[1,2], Marie-Pierre Buisine[9,10], Guillaume Piessen[9,11], Ada Collura[1], Jean-François Fléjou[1,4], Aurélien de Reynies[5], Florence Coulet[1,2]*, François Ghiringhelli[12,13]*, Thierry André[1,3]*, Vincent Jonchère[1,2]*, Alex Duval[1,2]*£

1) Sorbonne Université, INSERM, Unité Mixte de Recherche Scientifique 938 and SIRIC CURAMUS, Centre de Recherche Saint-Antoine, Equipe Instabilité des Microsatellites et Cancer, Equipe labellisée par la Ligue Nationale contre le Cancer, F-75012 Paris, France

2) Sorbonne Université, Genetics Department, AP-HP.Sorbonne Université, hospital Pitié-Salpêtrière, F-75012 Paris, France

3) Sorbonne Université, department of medical oncology, AP-HP, hôpital Saint-Antoine, F-75012 Paris, France

4) Sorbonne Université, department of pathology, AP-HP, hôpital Saint-Antoine, F-75012 Paris, France

5) Programme “Cartes d'Identité des Tumeurs”, Ligue Nationale Contre le Cancer, Paris, France

6) Service de Neurologie 2-Mazarin, Sorbonne Université, Inserm, CNRS, UMR S 1127, Institut du Cerveau, ICM, AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière - Charles Foix, 47-83 boulevard de l'Hôpital, 75013, Paris, France

7) AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière - Charles Foix, Service de Neurologie 2-Mazarin, F-75013, Paris, France

8) AP-HP, Hôpitaux Universitaires La Pitié Salpêtrière - Charles Foix, Département de Neuropathologie, F-75013, Paris, France

9) Univ. Lille, CNRS, Inserm, CHU Lille, UMR9020-U1277 CANTHER - Cancer Heterogeneity Plasticity and Resistance to Therapies, F-59000 Lille, France

10) Molecular Oncogenetics, Department of Biochemistry and Molecular Biology, Lille University Hospital, F-59000 Lille, France

11) University Lille, Department of Digestive and Oncological Surgery, Claude Huriez University Hospital, F-59000 Lille, France

12) Department of Medical Oncology, Centre Georges-François Leclerc, Dijon, France

13) INSERM, UMR866, Burgundy University

*** Co-leadership**

Correspondence should be addressed to:

Alex Duval, alex.duval@inserm.fr

Grant support: This work was supported by grants from SIRIC CURAMUS and the LIGUE NATIONALE CONTRE LE CANCER.

Disclosures: The authors have declared that no conflicts of interest exist.

Author Contributions: study concept and design: TR, VJ, FR, AC, AD; acquisition of data: TR, RC, MSv, QL, OB, AS, AC, EG, PC, PB, MPB, GP, FC, RN, MA, MSa, MT, FB, FR, PLR, JFF, ADR, FG, TA, VJ, AD; analysis and interpretation of data: TR, RC, FR, VJ, AD; drafting of the manuscript: TR, FR, VJ, AD; critical revision of the manuscript for important intellectual content: RC, MSv, QL, OB, AS, AC, EG, PC, FC, RN, MA, MSa, MT, CD, FB, PLR, JFF, ADR, FG, TA; obtained funding; AD; technical support: TR, MSv, QL, OB, AS, EG, PB, MPB, PC, FC, MA, FB, JFF, VJ; study supervision: AD.

Aknowledgments: We thank Dr Catherine Vermaut and Lucie Delattre (Lille university hospital) for their help with acquisition and interpretation of MSI data.

ABSTRACT

Background & Aims: Next generation sequencing (NGS) was recently approved by the FDA to detect microsatellite instability (MSI) arising from defective mismatch repair (dMMR) in patients with metastatic colorectal cancer (mCRC) prior to treatment with immune checkpoint inhibitors (ICI). In this study, we aimed to evaluate and improve the performance of NGS to identify MSI in CRC, especially dMMR mCRC treated with ICI.

Methods: CRC samples used in this post-hoc study were reassessed centrally for MSI and dMMR status using the reference methods of pentaplex PCR and immunohistochemistry (IHC). Whole exome (WES) was used to evaluate MSISensor, the FDA-approved and NGS-based method for assessment of MSI. This was performed in (i) a prospective, multicenter cohort (C1) of 102 mCRC patients (25 dMMR/MSI, 24 treated with ICI) from clinical trials NCT02840604 and NCT033501260, (ii) an independent retrospective, multicenter cohort of 113 patients (C2, 25 mCRC, 88 non-mCRC, all dMMR/MSI untreated with ICI), (iii) and a publicly available series of 118 CRC patients from the TCGA (C3, 51 dMMR/MSI). A new NGS-based algorithm, namely MSICare, was developed. Its performance for assessment of MSI was compared to MSISensor in C1, C2 and C3 at the exome-level or after downsampling sequencing data to the MSK-Impact™ gene panel. MSICare was validated in an additional retrospective, multicenter cohort (C4) of 152 new CRC patients (137 dMMR/MSI) enriched in MSH6 and PMS2 deficient tumors (35 dMSH6, 9 dPMS2) following targeted sequencing of samples with an optimized set of microsatellite markers (MSIDIAG).

Results: At the exome-level, MSISensor was highly specific but failed to diagnose MSI in 16% of MSI/dMMR mCRC from C1 (4/25; sensitivity 84%, 95%CI: 63.9%-95.5%), 32% of mCRC (8/25; sensitivity 68%, 95%CI: 46.5%-85.1%) and 9.1% of nmCRC from C2 (8/88;

sensitivity 90.9%, 95%CI: 82.9%-96%), and 9.8% of CRC from C3 (5/51; sensitivity 90.2%, 95%CI: 78.6%-96.7%). Misdiagnosis included 4 mCRCs treated with ICI of which 3 showed an overall response rate without progression at this date. At the exome-level, reevaluation of the MSI genomic signal using MSICare detected 100% of cases with true MSI status amongst C1 and C2. Further validation of MSICare was obtained in CRC tumors from C3, with 96.1% concordance for MSI status. Whereas misdiagnosis with MSISensor even increased when analyzing downsampled WES data from C1 and C2 with microsatellite markers restricted to the MSK-Impact gene panel (sensitivity 72.5%, 95%CI: 64.2-79.7%), particularly in MSH6 deficient setting, MSICare sensitivity and specificity remained optimal (100%). Similar results were obtained with MSICare following targeted NGS of tumors from C4 with the optimized microsatellite panel MSIDIAG (sensitivity 99.3%, 95%CI: 96%-100%; specificity 100%).

Conclusions: In contrast to MSISensor, the new MSICare test we propose performs at least as efficiently as the reference method, MSI PCR, to detect MSI in CRC regardless of the defective MMR protein under both WES and targeted NGS conditions. We suggest MSICare may become rapidly a reference method for NGS-based testing of MSI in CRC, especially in mCRC where accurate MSI status is required before the prescription of ICI.

Key words: Microsatellite instability (MSI) / Mismatch repair deficiency (dMMR), Next-generation sequencing, Diagnostic test, Reference methods, Immunotherapy

INTRODUCTION

The human tumor phenotype referred to as microsatellite instability (MSI) is associated with inactivating alterations in mismatch repair (MMR) genes. MSI was first reported in inherited tumors associated with Lynch syndrome. This is one of the most frequent cancer predisposition syndromes in humans and requires specific care and genetic counseling. MSI was later observed in sporadic colorectal cancer (CRC) and more rarely in other primary tumors [1-4]. Tumors with MSI generally show a dense infiltration with cytotoxic T-cell lymphocytes [5]. Recently, it was reported that MSI tumors and notably MSI CRC resist this hostile immune microenvironment by overexpressing immune checkpoint (ICK)-related proteins to allow immune-escape [6, 7]. Furthermore, MSI status was shown by our team and others to predict clinical benefit from ICK inhibitors (ICI) in patients with metastatic CRC (mCRC) [8-12]. These observations have led to international guidelines recommending universal MSI/dMMR screening of all newly diagnosed CRC [13]. There is also increasing evidence to support the evaluation of MSI status in all human tumors, regardless of the primary tissue of origin.

Several specialized cancer centers, including ours, have aimed to standardize and validate the accepted reference methods for testing MSI and dMMR in CRC, *i.e.* polymerase chain reaction (PCR)-based methods for MSI [14-16] and immunohistochemistry (IHC) for dMMR (see also [17] for review). In mCRC, we recently highlighted that misinterpretation of the results for MSI and/or MMR testing using these gold standard methods could account for most cases showing primary resistance to ICI [18]. In the meantime, an alternative FDA-approved method based on next generation sequencing (NGS) technology was reported for MSI screening in

pan-cancer, including CRC [19]. This was based on the use of an algorithm, namely MSISensor, that analyzes sequence reads at designated microsatellite regions in tumor and paired normal samples and reports the percentage of unstable loci as a cumulative score in the tumor [20]. However, the diagnostic performance of MSISensor has yet to be evaluated in patient cohorts where the MSI/dMMR status has already been established using the reference IHC and MSI-PCR methods, notably in the prospective setting of mCRC patients treated with ICI.

The aim of the present study was therefore to evaluate the performance of MSISensor for the detection of MSI in dMMR/MSI CRC, especially dMMR mCRC treated with ICI. CRC samples used in this post-hoc analysis were centrally reassessed for MSI and dMMR status using the reference methods of pentaplex PCR and immunohistochemistry (IHC). We analyzed samples from multicenter, prospective series of CRC patients involved in clinical trials with ICI (NCT02840604 and NCT033501260) and two large retrospective, multicenter independent series of mCRC and non-metastatic CRC (nmCRC), as well as a publicly available series of CRC (TCGA). Importantly, cohorts were enriched in MSH6-deficient tumors in which MSI is known to be more difficult to detect [21] and in PMS2-deficient tumors for which limited data are available. Overall, our results demonstrate that the FDA-approved NGS-based diagnostic test for identifying MSI in mCRC and nmCRC gave inaccurate results when compared with the gold standard reference methods. This misdiagnosis included patients that showed a positive response to ICI but would not have been treated if MSISensor alone had been used for MSI screening without reference to the IHC and MSI PCR methods. Next, full or partial WES data restricted to the MSK-Impact™ gene panel were exploited to improve detection of the MSI genomic signal in CRC. This enabled us to design and validate a newly optimized algorithm, namely MSICare. The high accuracy of MSICare for the detection of MSI in CRC was validated under both analysis of full or partial WES data or following targeted NGS of tumors with an optimized panel of microsatellite markers. These

results should allow MSICare to become a future reference test for assessing MSI in colorectal tumors and putatively in non-colorectal dMMR tumors, especially for dMMR metastatic patients prior to therapeutic decision.

MATERIALS AND METHODS

Study populations

The clinical rationale and design of this study are presented in **Figure 1**. The origin of patients enrolled in the study is shown in **Supplementary Table S1**. One hundred and two patients with mCRC (Cohort C1, **Fig. 1**) originated from two multicenter French clinical trials (NCT02840604 and NCT033501260) which accrued patients between May 2015 and November 2019. NCT02840604 aimed to show that exome analysis is feasible in the routine care of patients, thereby improving access to targeted therapies and improving the detection of genetic cancer predisposition. Genomic sequencing (WES) was performed at the Georges-Francois Leclerc Cancer Center, Genomic and Immunotherapy Medical Institute, Dijon, France. Patients were eligible if they presented with a locally advanced, non-operable or metastatic cancer that had progressed during at least one line of systemic therapy. The NIPICOL trial (NCT033501260) involves treatment of MSI/dMMR mCRC patients with nivolumab (anti-PD-1) and ipilimumab (anti-CTLA-4). mCRC response to ICI was determined according to Response Evaluation Criteria in Solid Tumors (RECIST) [22]. Twenty-six cases from the NCT033501260 cohort were treated with ICI. Of these, 23 were confirmed to be MSI/dMMR and 3 were identified later as MSS/pMMR following reassessment of their MSI and MMR status centrally. Whole Exome sequencing (WES) was performed by IntegraGen SA (Evry, France). All patients provided signed informed consent for the trials and genomic analysis. After giving consent, patients underwent consultation with a geneticist to explain the consequences of constitutional genetic testing. Following this consultation, the

patient could accept or refuse to provide a blood sample for constitutional exome analysis. This trial protocol was approved by an institutional review committee and performed in accordance with the Declaration of Helsinki.

A historical retrospective cohort was also analyzed by WES in the same conditions (cohort C2, **Fig. 1**). This comprised 25 patients with mCRC that were diagnosed between 1998 and 2016 in 6 French hospitals as MSI or dMMR [18], as well as 88 patients from the Saint Antoine Hospital, Paris, who were diagnosed between 1998 and 2007 as having MSI/dMMR nmCRC [23].

We further evaluated and compared the performance of MSICare and MSISensor using WES data in a third independent tumor cohort (C3). This included 118 CRC patients whose MSI status was previously assessed by PCR using the Bethesda microsatellite panel and whose WES data was publicly available from the TCGA. All CRC patients with MSI-H (*N* = 51, MSI-High) or MSI-L (*N* = 14, MSI-Low) and a similar proportion of patients with MSS (*N* = 53) were included [24]. C3 also included 382 extra-colonic tumors from the TCGA with a relatively high incidence of MSI, *i.e.* gastric (53 MSI-H, 9 MSI-L, 42 MSS) and endometrial (159 MSI-H, 17 MSI-L, 102 MSS) cancers.

Finally, a last retrospective cohort was examined (cohort C4, Fig. 1) using targeted NGS because WES is not routinely used in clinical care (see below for details). C4 was retrospective, non-consecutive, assembling 152 new patients from the Saint Antoine Hospital, Paris, and the Lille University Hospital, who were previously diagnosed as having MSI/dMMR or MSS/pMMR CRC (137 MSI, 15 MSS) using MSI PCR and IHC. dMMR/MSI CRC cases from the Saint-Antoine Hospital were previously diagnosed as being dMMR/MSI between 1998 and 2021 regardless of the MMR defect detected in the tumor. Both tumor and non-tumor DNA material was available for these cases and they were not previously analyzed

by WES (no overlap with the C2 cohort). dMMR/MSI CRC cases from the Lille University Hospital with material available (biopsy or surgical resection) from 2016 to 2021 displayed isolated loss of MSH6 or PMS2 expression. They were selected to further evaluate the performance of MSICare for identifying MSI in these rare dMMR/MSI CRC settings, in particular for MSH6-deficient CRC which are known to be more difficult to diagnose [21].

All patients provided written consent and the study was approved by the institutional review boards/ethics committees of the participating centers.

Samples

In the prospective cohort (C1, clinical trials NCT02840604 and NCT033501260), the mCRC samples were formalin fixed and paraffin-embedded (FFPE) and were composed of either the primary or metastatic tumor tissue. In the C2 retrospective cohort, all nmCRC samples (N = 88) were stored at -80 °C until DNA extraction. For mCRC patients (N = 25), both the primary tumor and metastasis preserved in FFPE (N = 45; 25 primary tumors and 20 metastases) were collected and analyzed whenever possible in order to provide a more complete description of this rare CRC subtype. For the public retrospective cohort (C3), frozen tissue samples were collected from the primary tumor sites (colorectum, stomach, endometrium) as described [24]. In the C4 retrospective cohort, CRC samples (N = 152; primary or metastases) and matched non tumor samples were either FFPE (N = 87) or frozen (N = 65) in order to appreciate the feasibility of the MSICare method under various technical conditions and qualities of DNAs. Matched normal colonic mucosa samples were considered in all cohorts to perform NGS-based MSI screening.

Immunohistochemistry and MSI-PCR

All CRC samples from C1, C2 and C4 used in this post-hoc study were centrally

reassessed in expert centers involved in this study (Saint-Antoine hospital, Paris France, Georges-Francois Leclerc Cancer Center, Dijon, France and Lille University Hospital, France) for dMMR status using immunohistochemistry (IHC) and for MSI using polymerase chain reaction (PCR) as previously described [14-16].

MSISensor End Points

False negatives were defined as samples initially diagnosed as MSI and/or dMMR using MSI-PCR and IHC, respectively, but showing a negative MSISensor Score (≤ 10%) when considering (i) the full exome data or downsampled exome data restricted to the MSK-Impact™ gene panel (C1, C2, C3)[19, 20]; (ii) or MSIDIAG microsatellite panel of markers (See below) following targeted sequencing of tumors (C4). This was done by central assessment at the Saint-Antoine Hospital, Paris (C1, C2, C3, C4), at the Georges-Francois Leclerc Cancer Center, Dijon (C1), or at the Lille University Hospital (C4). The sensitivity of MSISensor was calculated as the percentage of true-positive cases amongst the total of true-positive and false-negative cases.

Whole exome sequencing and NGS-based MSI diagnosis with MSISensor

For the prospective (C1) and retrospective (C2) cohorts, the WES procedure was performed as recommended by the manufacturer (SureSelect Human Exon Kit v5, 75 MB; Agilent, Les Ulis, France) and as previously described [25]. For metastatic tumor samples (C1 and C2), the generated reads were mapped to the reference genome hg38 (GRCh38), while for the retrospective non-metastatic samples (C2), the reads were mapped onto hg19 (GRC37). Sequencing data were comprised approximately between 50X depth (normal samples) and 200X depth (tumor samples). MSISensor was used at the default setting to evaluate the mutation status of microsatellites from the WES data [20].

Implementation of the optimized NGS-based MSICare method to increase the sensitivity of MSI detection in CRC and in other tumors

A new method (MSICare) was developed to optimize the detection of MSI based on comparison of the read distribution between normal and tumor samples from full WES data (C1, C2 and C3). Mononucleotide repeats (MNR) with a length ≥ 12 base pairs (bp) were considered for analysis only if they were covered by at least 20 mapping reads in both normal and tumor samples. The total number of reads covering each candidate MNR was then normalized (arbitrary value of 100) in tumor and matched healthy tissue. For each MNR and each deletion size, the normalized number of reads in healthy tissue was subtracted from the normalized number of reads in tumor tissue [ΔRatio = %Tumor-%Normal] to generate an MSI signal corresponding to the sum of ΔRatio values for all candidate MNR. The ΔRatio value was then adjusted by estimating the tumor purity (TP) for each tumor sample, with the estimated TP corresponding to the median value of the MSI signal for all MNR with a length ≥ 14 bp covered by at least 30 reads in tumor and 20 reads in normal tissue. The adjusted value for ΔRatio was then used to classify a given MNR as wild type (ΔRatio-adjusted = ΔRatio x Estimated TP < 50%) or mutated (ΔRatio-adjusted = ΔRatio x Estimated TP ≥ 50%) given that observed microsatellite mutations can be either heterozygous or homozygous in primary tumor samples. Finally, the MSICare score for tumor samples corresponds to the percentage of microsatellites that were mutated amongst the total number of microsatellites analyzed using this approach. The scripts and documentation are available through Github at https://github.com/CRSA-MSI/MSICare.

MSICare cutoff determination

A cutoff value for MSICare was estimated in order to optimize the differentiation of MSI from MSS samples in the C1 and C2 cohorts. This was done using the cutpointr package

(version 1.0.32), which estimates optimal cutoff points in binary classification tasks and validates their performance using bootstrapping. A cutoff point of 20 was determined using full WES data in a discovery set of 77 MSS and 138 MSI (C1 + C2; CRC, Discovery set) and then applied to a validation set of MSI (C3; CRC and non-CRC, Validation set) from public TCGA data (see the Results section for further details). The same cutoff was tested again to test MSICare for identifying MSI in the same cohorts of CRC patients when considering only partial WES data restricted to the MSK-Impact^{TM} gene panel.

Diagnosing MSI in CRC with MSICare following targeted sequencing of paired tumoral and normal mucosa samples with an optimized panel of microsatellite markers

From the viewpoint of clinical application, MSI test is important not only in Whole exome sequencing, but also in panel testing. The performance of MSICare as compared to MSISensor was assessed again in the additional independent, multicenter CRC cohort (C4) using the same cutoff. Sequencing of this cohort on paired tumor and normal mucosa samples was performed using an optimized targeted panel of microsatellite markers, namely MSIDIAG. This panel includes 441 mononucleotide repeats which have been selected among the MNR harboring a size of 12 bp or more whose instability was exclusively observed in MSI tumor samples from C1, C2 and C3 following WES (low frequency of somatic mutations in MSS CRC; chi-squared test with p-value < 0.05). After capture and sequencing, reads were mapped to the Human genome build (hg38) with depth of coverage comprised between 100X and 500X. The diagnosis of MSI was assessed using MSISensor or MSICare procedure exactly as this was performed previously from WES data in C1, C2 and C3 (see above).

RESULTS

Frequent misdiagnosis of MSI with MSISensor in both mCRC and nmCRC

All CRC samples from C1 and C2 were centrally reassessed for MSI and dMMR status using the gold standard reference methods of pentaplex PCR and IHC (**Fig.1, Supplementary Table S1,** and data not shown). MSISensor confirmed the status of 77 MSS/pMMR mCRC from the prospective C1 cohort (**Fig.1** and **Fig. 2A**; MSISensor score ≤ 10%). However, it failed to confirm the status of 4 of the 25 MSI/dMMR mCRC samples (**Fig. 2A**; MSISensor score ≤ 10%). The frequency of misdiagnosis in C1 was therefore 16% (*N* = 4/25; sensitivity 84%, 95%CI: 63.9%-95.5 %).

The sensitivity of MSISensor was further assessed in 25 mCRC patients with MSI/dMMR from the retrospective C2 cohort (**Fig. 1**). In mCRC, the frequency of misdiagnosis was even higher at 32% (*N* = 8/25; sensitivity 68%, 95%CI: 46.5%-85.1%) (**Fig. 2B**). **Supplementary Figure S1** shows the performance of MSISensor according to the metastatic site for these 25 mCRC with MSI/dMMR. In 88 nmCRC patients with MSI/dMMR from the C2 cohort, misdiagnosis occurred in 9.1% (*N* = 8/88, 9%; sensitivity 90.9%, 95%CI: 82.9%-96%) (**Fig. 2B**). The sensitivity of MSISensor was finally assessed in the public C3 cohort of CRC patients that included both nmCRC and mCRC (**Fig. 1**). The frequency of missed diagnoses was again very similar at 9.8% (*N* = 5/51), giving a sensitivity of 90.2% (95%CI: 78.6%-96.7%) in patients with MSI/dMMR CRC. This included one misdiagnosed case of mCRC (1/3, 33%) (**Fig. 2C**). MSISensor confirmed the status of all but 2 MSS/pMMR mCRC from C3, thus indicating the major limitation of this method was its lack of sensitivity. In **Figure 2D**, we analyze the performance of MSISensor taking into account the nature of dMMR defect in tumors. The results indicate that the sensitivity of this test is likely to be especially low in MSH6-deficient or PMS2-deficient CRC (sensitivity 71.4%, 95%CI: 29%-96.3%), as

expected [21]. The overall performance of MSISensor in the C1 cohort compared to the C2 and C3 cohorts is shown in **Table 1.**

Identifying weaknesses and limits of MSISensor by deciphering the MSI genomic signal of DNA repeats in CRC

A density plot was created to show fluctuations in the MSISensor score for the C1, C2 and C3 patient cohorts analyzed in this study (**Fig. 3A**). The MSI/dMMR and MSS/pMMR status of all samples in cohorts C1 and C2 were pooled as these had previously been validated centrally. CRC samples from the public cohort C3 were considered separately since we were unable to independently confirm the status of these tumors using IHC and MSI-PCR. The density profiles clearly highlight the lack of sensitivity of MSISensor for the detection of MSI in dMMR CRC, as already shown above for the 3 cohorts (**Fig. 2** and **Table 1**).

We next hypothesized that it should be possible to improve the detection of MSI in CRC by modifying certain parameters in the analysis of NGS data. In support of this, WES analyses revealed that MSISensor lacked sensitivity because: (i) MNR sequences were by far the most unstable category of microsatellites in dMMR colon tumors and therefore better at distinguishing MSI from MSS CRC than other types of repeat used by MSISensor (*e.g.* di-, tri-, tetra-, penta-) (**Supplementary Figure S2A**); (ii) the ability of MNR to discriminate between MSS and MSI colon tumors was dependent upon their length, with long MNR of ≥ 12 bp found to be the most discriminating compared to other microsatellites used by MSISensor (**Supplementary Figure S2B**); (iii) MSISensor was not suitable for detecting MSI in CRC samples with an estimated TP of less than 30-40%. This is an important limitation to the sensitivity of MSISensor in primary MSI CRC due to the often high levels of contamination with non-tumor and inflammatory pMMR/MSS cells (**Supplementary Figure**

S3; see also our review [26] and original publications for further details [15, 16, 23, 25, 27]). WES analyses also revealed that MSISensor lacks specificity for two reasons. Firstly, the MSISensor computational tool confused the true MSI signal with allelic losses (LOH) for some of the MNR. LOH occurs frequently in MSS colon tumors with high levels of chromosomal instability (**Supplementary Figure S4A**). Secondly, stuttering by DNA polymerase during the PCR reaction occurs frequently at microsatellites and in particular at long MNR. A misdiagnosis of MSI can therefore occur when small 1 bp deletions in these microsatellites are considered by MSISensor to represent MSI (**Supplementary Figure S4B**). **Figure 3B** shows a summary of the various pitfalls that have been identified for MSISensor, leading us to develop a more refined NGS-based bioinformatic tool, namely MSICare.

Designing and validating MSICare to improve NGS-based detection of MSI in CRC

To avoid the abovementioned pitfalls of MSISensor, we next designed a new computational tool referred to as MSICare to accurately detect MSI in CRC from C1 and C2, based on analysis of their WES profile. In contrast to MSISensor, MSICare identifies true MSI signals defined as somatic deletions of at least 2 bp in length that occur in long MNR (≥ 12 bp) in DNA from dMMR cancers but not in DNA from paired normal tissue (see **Fig. 3B** and Materials and Methods for further details). A Receiver Operating Characteristic (ROC) curve was constructed assuming binary classification of the MSICare score. This revealed a perfect discrimination between dMMR and pMMR CRC in C1 and C2, with 100% sensitivity and 100% specificity when using a cut-off value of 20% (**Fig. 4A**). dMMR/MSI samples showed a mean MSICare score above 80% with little dispersion around this value, whereas the mean MSICare score of pMMR/MSS samples was below 10%. MSICare thus appears to be very effective in discriminating MSI from MSS CRC cases. The high level of discrimination achieved using the cut-off value of 20% was validated in the public C3 cohort

(**Fig. 4B** right panel) and led us to correct 3 cases amongst 5 that were true MSI that showed false negative status with MSISensor. Of interest, the two remaining CRC samples with negative MSISensor status that were previously classified as MSI by PCR according to TCGA remained unambiguously MSS with MSICare. Detailed analysis of the exomic profile of both these tumors revealed very few mutations in microsatellites located within coding regions of known MSI target genes (**Supplementary Figure S5**. Furthermore, one of the samples showed pMMR according to TCGA, thus suggesting an equivocal MSI status. The overall performance of MSICare in the C1 cohort compared to the C2 and C3 cohorts is shown in **Table 2.** In **Figure 4C**, we analyze the performance of MSICare taking into account the nature of dMMR defect in tumors. The results indicate that the sensitivity of this test remain optimal in MSH6-deficient or PMS2-deficient colon tumors from this cohort (sensitivity 100%).

Confirmation of the performance of MSICare in detecting MSI in CRC using targeted NGS

Because WES is not routinely used in clinical care, we finally aimed to confirm the high performance of MSICare for detecting MSI following targeted sequencing of CRC samples as compared to paired normal mucosa samples. This was first done in C1 and C2, considering only the microsatellites included in the restricted MSK-IMPACT^{TM} gene panel (see Materials and Methods for details). The overall number of false negative cases detected amongst MSI/dMMR from these cohorts under these conditions was important with MSISensor for all 3 versions of MSK-IMPACT^{TM} (**Fig. 5A** and **Supplementary Figure S6**), particularly in MSH6 and PMS2 deficient settings (sensitivity 28.6%, 95%CI: 4.9%-62%) (**Fig. 5A**). By, contrast, the performance of MSICare remained optimal under the same conditions (sensitivity 100%) (**Fig.5B**).

Next, we generated an optimally designed panel of 441 mononucleotide repeats (length≥ 12 pb and unstable in MSI tumors; See Methods for Details) called MSIDIAG. Using this panel we confirmed that, in the C4 cohort including 152 patients (137 MSI, 15 MSS) and that was enriched in CRC with MSH6 (35 patients) or PMS2 (9 patients) deficiency, MSICare still optimally detected MSI in CRC regardless of MMR defect (Sensitivity 99.3%, 95%CI: 97.8%-100.7% ; Specificity 100%) (**Fig. 5C**) whereas MSISensor remained less sensitive while becoming unspecific expectedly (Sensitivity 97.1%, 95%CI: 92.7%-99.2%; Specificity 73.3%, 95%CI: 44.9%-92.2%) (**Supplementary Figure S7**).

DISCUSSION

Several publications have recently highlighted the potential of NGS for the detection of MSI in human cancers by using distinct computational algorithms [19, 20, 28-31]. Amongst these, MSISensor has received FDA approval and is used to guide the prescription of ICI therapy in patients with metastatic cancer, regardless of the primary location of the tumor. MSISensor has been tested on advanced solid cancers including a large number of CRC. However, the performance of this NGS-based test has yet to be evaluated in a large series of CRC previously assessed for MSI/dMMR status using the reference PCR and IHC methods. The accuracy of MSISensor is especially important for patients deemed as MSI/dMMR mCRC and subsequently treated with ICI, but also more generally in CRC regardless of tumor staging to help the detection of Lynch syndrome. It's also important in nmCRC because adjuvant chemotherapy with fluoropyrimidines may be ineffective or even detrimental in MSI/dMMR CRC patients with localized diseases, especially in those with stage II disease [32]. In this study we provide clear evidence that MSISensor lacks sensitivity for the detection of

MSI. This was shown in large cohorts of mCRC and nmCRC samples that were previously confirmed as MSI/dMMR or MSS/pMMR by IHC and MSI-PCR methods performed in large, specialized test centers. These results are of particular clinical relevance for ICI therapy. They highlight that in a prospective cohort of MSI mCRC patients, the consideration of results from MSISensor alone in the absence of MSI-PCR and IHC testing would have led to significant amount of cases not being offered ICI treatment. Of interest, of the 4 MSI CRC patients not detected by MSISensor, 3 were found to be responsive to ICI treatment. The present results in CRC patients are consistent with those of another study that found NGS was unable to detect MSI in dMMR tumors from two prostate cancer patients who displayed prolonged positive response to ICK blockade therapy [33]. Both tumors showed a high mutational burden and a high density of intratumoral infiltration with CD3 cells. We therefore extrapolate that the low sensitivity of MSISensor for the detection of MSI is likely to apply to all tumor types, as suggested also by the analysis of gastric and endometrial tumors from the TCGA in the present study (see **Supplementary Figure S8**).

The new MSICare bioinformatic tool proposed here for the detection of MSI shows much better performance compared to MSISensor. As an expert center for the analysis of MSI in clinical oncology, we have optimized this bioanalytic tool so that MSI detection in tumor DNA is highly sensitive while remaining specific. It has 100% sensitivity and specificity compared to PCR-MSI in the CRC cohorts tested here, thus matching the performance of the gold standard IHC and MSI-PCR methods. The use of MSICare makes it possible to diagnose MSI in CRC that is highly contaminated with stromal tissue, which is frequently the case in MSI primary tumors. Of note, this new algorithm shows the same performance for both FFPE and frozen primary or metastatic tissue samples regardless of their primary MMR defect in MLH1, MSH2, MSH6 or PMS2. It was notably checked to be relevant, in contrast to MSISensor, for identifying MSI in MSH6-deficient CRC in which MSI is more difficult to

diagnose [21] and in PMS2-deficient CRC which are scarce. Importantly, and in contrast to MSISensor again, its performance for assessment of MSI remained optimal when tested with the full or partial exome data restricted to the MSK-Impact^{TM} panel of markers. MSICare also showed great performances when exploited with an optimally designed set of microsatellite markers following targeted sequencing of tumor samples (MSIDIAG). The outstanding diagnostic performance of MSICare to detect MSI with different sequencing strategies on independent series of CRC validates with a high level of evidence the relevance of this method to detect MSI in CRC. The MSIDIAG panel includes mononucleotide repeats that are of particular interest for detecting MSI in tumor DNA and it is therefore recommended to use this panel with MSICare in targeted sequencing analyses for optimal sensitivity of this assay. In summary, these data establish that MSICare has the potential to become a new NGS-based international reference method for the determination of MSI phenotype in CRC from WES or targeted NGS using home-made or FDA-approved panels. It should become very useful for translational research, clinical trials and in routine clinical practice in the management of CRC patients, especially as MSI is becoming an indispensable theranostic biomarker in the metastatic setting.

There were several limitations to this study. Firstly, although our results suggest the performance of MSICare in a public series of GC and EC from the TCGA (see **Supplementary Figure S8**), further evaluation in non-colorectal cancer types is required. Secondly, we investigated a relatively small prospective series of true MSI/dMMR mCRC patients treated with ICI. Larger prospective studies are required to confirm these findings in mCRC settings involving ICI treatment. In the future, we plan to further validate the exceptional performance of MSICare in CRC and non-CRC patients using the home-made MSIDIAG or commercialized panel of markers (*e.g.* MSK, Fondation One), especially for the metastatic setting with immunotherapy. This should allow further confirmation of MSICare as

an accurate NGS-based tool for the detection of MSI in human cancers. This comes at a critical time when MSI is increasingly becoming a theranostic marker to guide the therapy of patients at all stages of cancer and for a growing number of tumor types.

REFERENCES

1. Aaltonen LA, Peltomaki P, Leach FS, et al. Clues to the pathogenesis of familial colorectal cancer. Science 1993;260:812-6.
2. Fishel R, Lescoe MK, Rao MR, et al. The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell 1993;75:1027-38.
3. Thibodeau SN, Bren G, Schaid D. Microsatellite instability in cancer of the proximal colon. Science 1993;260:816-9.
4. Ionov Y, Peinado MA, Malkhosyan S, et al. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. Nature 1993;363:558-61.
5. Lothe RA, Peltomaki P, Meling GI, et al. Genomic instability in colorectal cancer: relationship to clinicopathological variables and family history. Cancer Res 1993;53:5849-52.
6. Llosa NJ, Cruise M, Tam A, et al. The vigorous immune microenvironment of microsatellite instable colon cancer is balanced by multiple counter-inhibitory checkpoints. Cancer Discov 2015;5:43-51.
7. Marisa L, Svrcek M, Collura A, et al. The Balance Between Cytotoxic T-cell Lymphocytes and Immune Checkpoint Expression in the Prognosis of Colon Tumors. J Natl Cancer Inst 2018;110.
8. Le DT, Uram JN, Wang H, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. N Engl J Med 2015;372:2509-20.
9. Overman MJ, Lonardi S, Wong KYM, et al. Durable Clinical Benefit With Nivolumab Plus Ipilimumab in DNA Mismatch Repair-Deficient/Microsatellite Instability-High Metastatic Colorectal Cancer. J Clin Oncol 2018;36:773-779.
10. Overman MJ, McDermott R, Leach JL, et al. Nivolumab in patients with metastatic DNA mismatch repair-deficient or microsatellite instability-high colorectal cancer (CheckMate 142): an open-label, multicentre, phase 2 study. Lancet Oncol 2017;18:1182-1191.
11. Le DT, Kim TW, Van Cutsem E, et al. Phase II Open-Label Study of Pembrolizumab in Treatment-Refractory, Microsatellite Instability-High/Mismatch Repair-Deficient Metastatic Colorectal Cancer: KEYNOTE-164. J Clin Oncol 2020;38:11-19.
12. Andre T, Shiu KK, Kim TW, et al. Pembrolizumab in Microsatellite-Instability-High Advanced Colorectal Cancer. N Engl J Med 2020;383:2207-2218.
13. Luchini C, Bibeau F, Ligtenberg MJL, et al. ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. Ann Oncol 2019;30:1232-1243.
14. Suraweera N, Duval A, Reperant M, et al. Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. Gastroenterology 2002;123:1804-11.

15. Buhard O, Cattaneo F, Wong YF, et al. Multipopulation analysis of polymorphisms in five mononucleotide repeats used to determine the microsatellite instability status of human tumors. J Clin Oncol 2006;24:241-51.
16. Buhard O, Lagrange A, Guilloux A, et al. HSP110 T17 simplifies and improves the microsatellite instability testing in patients with colorectal cancer. J Med Genet 2016;53:377-84.
17. Svrcek M, Lascols O, Cohen R, et al. MSI/MMR-deficient tumor diagnosis: Which standard for screening and for diagnosis? Diagnostic modalities for the colon and other sites: Differences between tumors. Bull Cancer 2019;106:119-128.
18. Cohen R, Hain E, Buhard O, et al. Association of Primary Resistance to Immune Checkpoint Inhibitors in Metastatic Colorectal Cancer With Misdiagnosis of Microsatellite Instability or Mismatch Repair Deficiency Status. JAMA Oncol 2019;5:551-555.
19. Middha S, Zhang L, Nafa K, et al. Reliable Pan-Cancer Microsatellite Instability Assessment by Using Targeted Next-Generation Sequencing Data. JCO Precis Oncol 2017;2017.
20. Niu B, Ye K, Zhang Q, et al. MSIsensor: microsatellite instability detection using paired tumor-normal sequence data. Bioinformatics 2014;30:1015-6.
21. You JF, Buhard O, Ligtenberg MJ, et al. Tumours with loss of MSH6 expression are MSI-H when screened with a pentaplex of five mononucleotide repeats. Br J Cancer 2010;103:1840-5.
22. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer 2009;45:228-47.
23. Collura A, Lagrange A, Svrcek M, et al. Patients with colorectal tumors with microsatellite instability and large deletions in HSP110 T17 have improved response to 5-fluorouracil-based chemotherapy. Gastroenterology 2014;146:401-11 e1.
24. Hause RJ, Pritchard CC, Shendure J, et al. Classification and characterization of microsatellite instability across 18 cancer types. Nat Med 2016;22:1342-1350.
25. Jonchere V, Marisa L, Greene M, et al. Identification of Positively and Negatively Selected Driver Gene Mutations Associated With Colorectal Cancer With Microsatellite Instability. Cell Mol Gastroenterol Hepatol 2018;6:277-300.
26. Duval A, Hamelin R. Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability. Cancer Res 2002;62:2447-54.
27. Dorard C, de Thonel A, Collura A, et al. Expression of a mutant HSP110 sensitizes colorectal cancer cells to chemotherapy and improves disease prognosis. Nat Med 2011;17:1283-89.
28. Salipante SJ, Scroggins SM, Hampel HL, et al. Microsatellite instability detection by next generation sequencing. Clin Chem 2014;60:1192-9.
29. Kautto EA, Bonneville R, Miya J, et al. Performance evaluation for rapid detection of pan-cancer microsatellite instability with MANTIS. Oncotarget 2017;8:7452-7463.
30. Vanderwalde A, Spetzler D, Xiao N, et al. Microsatellite instability status determined by next-generation sequencing and compared with PD-L1 and tumor mutational burden in 11,348 patients. Cancer Med 2018;7:746-756.
31. Johansen AFB, Kassentoft CG, Knudsen M, et al. Validation of computational determination of microsatellite status using whole exome sequencing data from colorectal cancer patients. BMC Cancer 2019;19:971.
32. Sargent DJ, Marsoni S, Monges G, et al. Defective mismatch repair as a predictive marker for lack of efficacy of fluorouracil-based adjuvant therapy in colon cancer. J Clin Oncol 2010;28:3219-26.

33. Antonarakis ES, Piulats JM, Gross-Goupil M, et al. Pembrolizumab for Treatment-Refractory Metastatic Castration-Resistant Prostate Cancer: Multicohort, Open-Label Phase II KEYNOTE-199 Study. J Clin Oncol 2019:JCO1901638.

LEGENDS TO FIGURES

Figure 1. Background and Study Design. FDA, Food and Drug Administration; MSI, microsatellite instable; CRC, colorectal cancer; mCRC, metastatic colorectal cancer; nmCRC, non-metastatic colorectal cancer; WES, whole exome sequencing; ICI, immune checkpoint inhibitors; IHC, Immunohistochemistry.

Figure 2. Reassessment of MSI using MSISensor in prospective and retrospective cohorts of metastatic and non-metastatic CRC whose MSI/dMMR or MSS/pMMR status had been previously assessed with gold standard reference methods

A) Boxplots show the percentage of mutated microsatellites (MSISensor score) obtained from WES of 25 MSI (red) and 77 MSS (blue) patients with metastatic CRC from a prospective cohort (Cohort 1, C1).

B) Boxplots show the percentage of mutated microsatellites (MSISensor score) obtained from WES of 88 MSI patients with non-metastatic CRC (left) and 25 MSI patients with metastatic CRC (right) from a retrospective cohort (Cohort 2, C2).

C) Boxplots show the percentage of mutated microsatellites (MSISensor score) obtained from WES of 118 TCGA patients, including 51 MSI, 14 MSI-L and 53 MSS patients (Cohort 3, C3).

D) Boxplots show the percentage of mutated microsatellites (MSISensor score) obtained from WES of patients from the C1 and C2 cohorts with either MLH1, MSH2, MSH6 or PMS2 deficient CRC.

A cutoff MSISensor score of 10 (FDA recommendation) was used to discriminate MSS from MSI tumors (green dotted line). Non-metastatic samples are represented by a circle and metastatic samples by a diamond. Horizontal barplots indicate the percentage of true negative (TN), true positive (TP), false negative (FN) and false positive (FP) for each cohort.

Figure 3. Improving the computational detection of MSI in CRC by identifying the weaknesses and limits of MSISensor

A) Density plot of the MSISensor score in C1 + C2 (left) and C3 (right) cohorts. The cutoff MSISensor score of 10 (FDA recommendation) was used to separate MSS from MSI tumors (green dotted line). The adjacent histograms represent the distribution of tumor samples according to their MSISensor score.

B) Schematic representation of the improvement in the MSI status detection method using NGS (See also Results section and **Supplementary Figure S3** for details).

Figure 4. Testing the diagnostic performance of MSICare

A) Density plot of the MSICare score in the C1 and C2 cohorts.

B) Density plot of the MSICare score in the C3 cohort.

C) Boxplots show the percentage of mutated microsatellites (MSICare score) obtained from WES of patients from the C1 and C2 cohorts with either MLH1, MSH2, MSH6 or PMS2 deficient CRC.

Horizontal barplots indicate the percentage of true negative (TN) and true positive (TP).

Figure 5. Comparative Performance of MSISensor and MSICare for the identification of MSI using targeted panel sequencing data

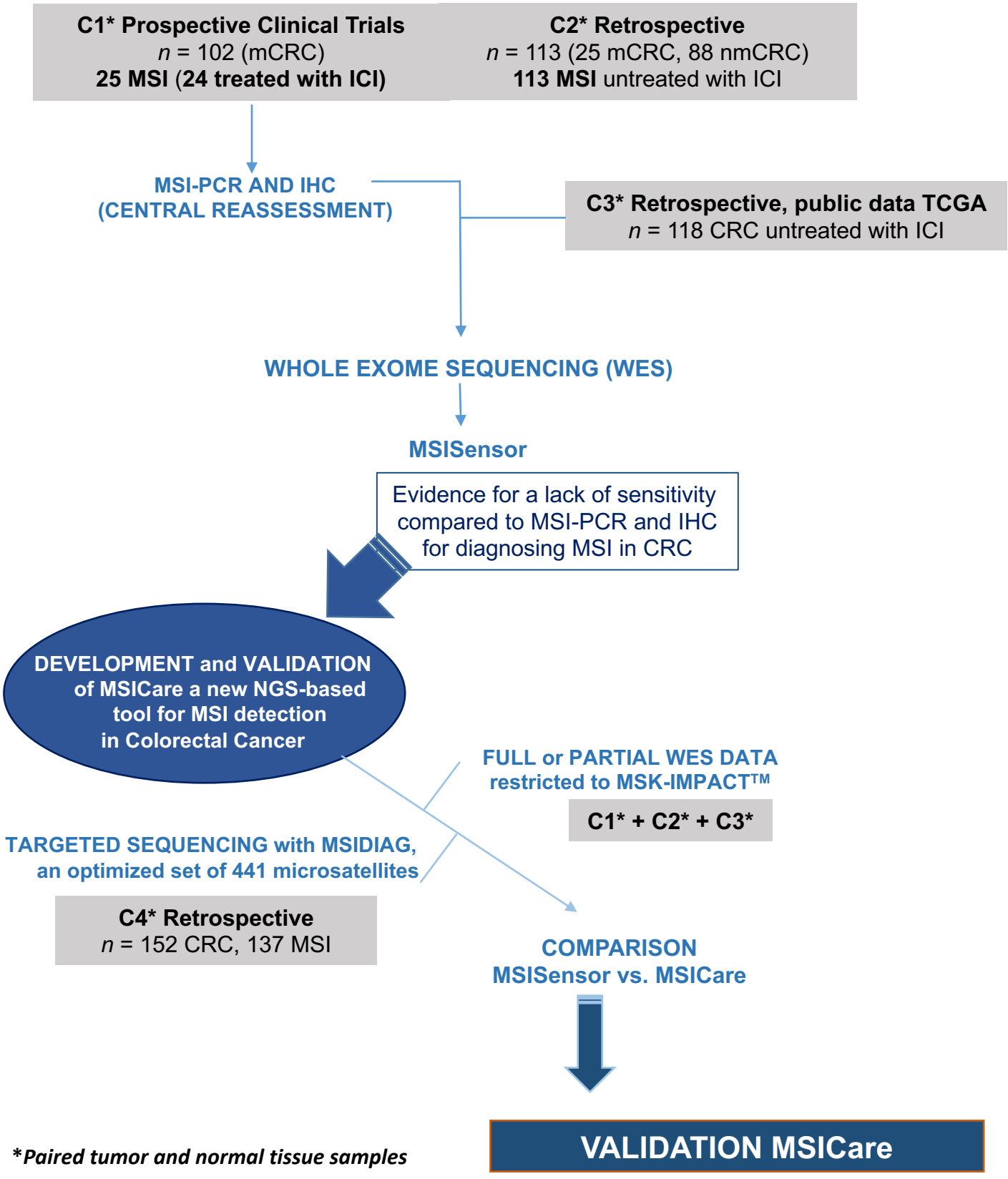
A) Boxplots show the percentage of mutated microsatellites (MSISensor score) obtained from the MSK-IMPACT for patients from the C1 and C2 cohorts (all patients, right panel; the same patients with either MLH1, MSH2, MSH6 or PMS2 deficient CRC, left panel). A cutoff MSISensor score of 10 (FDA recommendation) was used to discriminate MSS from MSI tumors (green dotted line).

B) Boxplots show the percentage of mutated microsatellites (MSICare score) obtained from the MSK-IMPACT^{TM} for patients from the C1 and C2 cohorts (all patients, right panel; the same patients with either MLH1, MSH2, MSH6 or PMS2 deficient CRC, left panel). A cutoff MSICare score of 20 was used to discriminate MSS from MSI tumors (green dotted line).

C) Boxplots show the percentage of mutated microsatellites (MSISensor score and/or MSICare score) obtained following targeted sequencing of patients from the C4 cohort (all patients, right panel; the same patients with either MLH1, MSH2, MSH6 or PMS2 deficient CRC, left panel). The pentaplex profile of the only one misdiagnosed case is shown in a box.

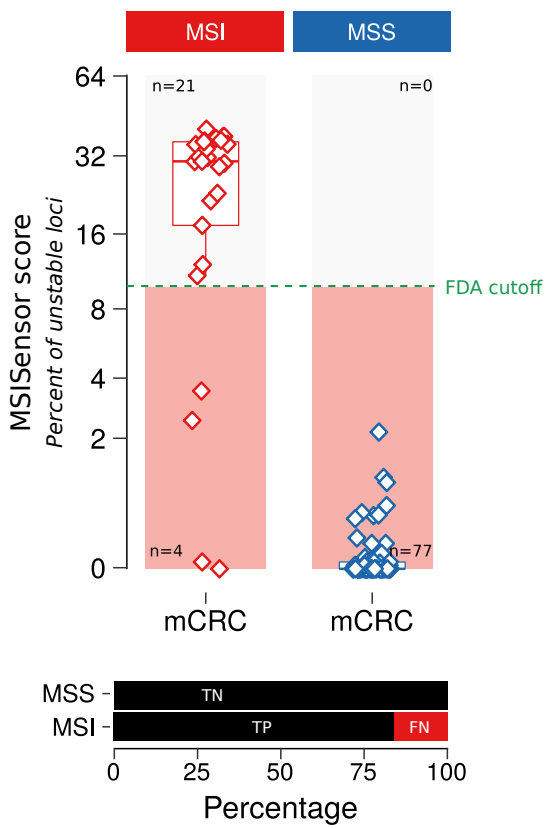
Horizontal barplots indicate the percentage of true negative (TN), true positive (TP), false negative (FN) and false positive (FP).

STUDY DESIGN

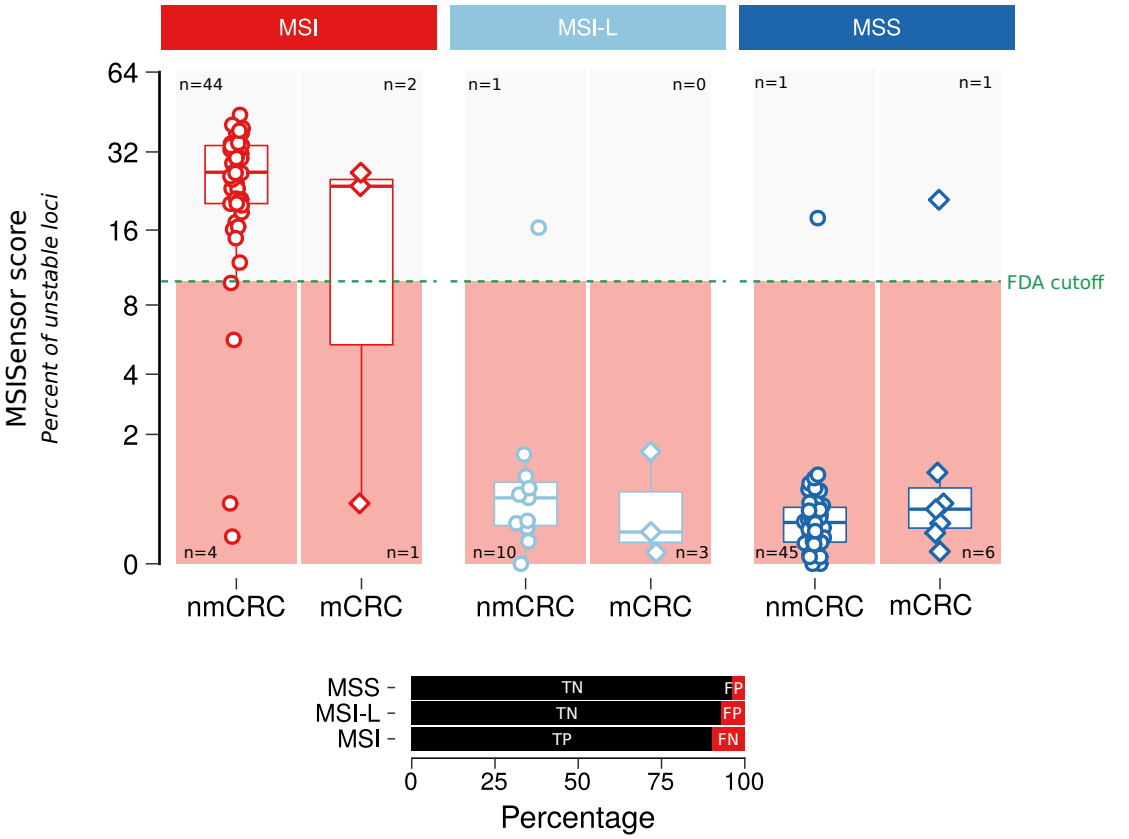


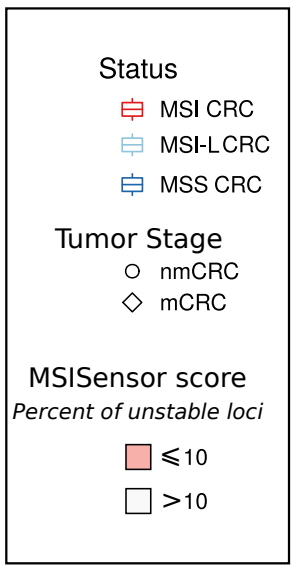
Paired tumor and normal tissue samples

A

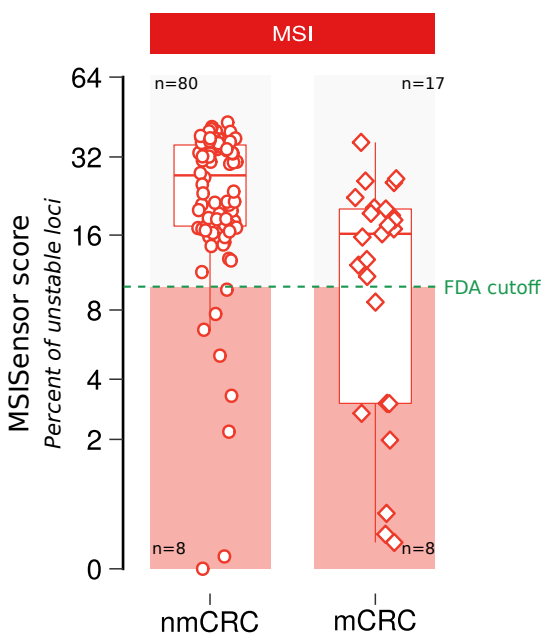

Prospective (C1, n=102)
mCRC
MSI
MSS
MSISensor score
Percent of unstable loci
64
32
16
8
4
2
0
n=21
n=0
n=4
n=77
FDA cutoff
mCRC
mCRC
MSS
MSI
TN
TP
FN
0
25
50
75
100
Percentage

C

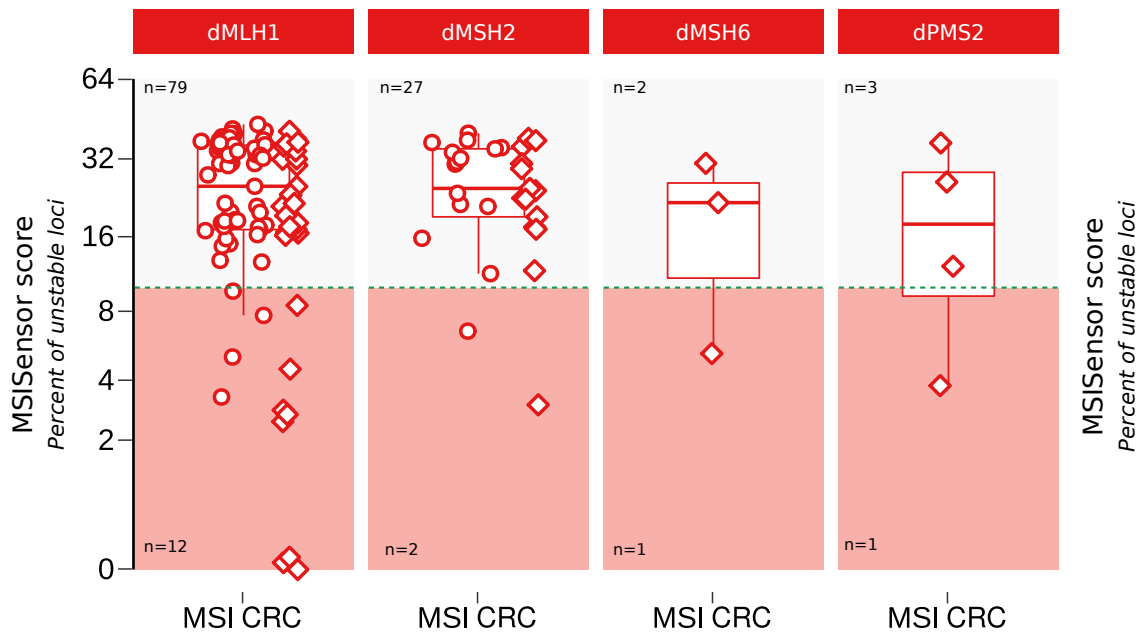

TCGA (C3, n= 118)
mCRC & nmCRC
MSI
MSI-L
MSS
MSISensor score
Percent of unstable loci
64
32
16
8
4
2
0
n=44
n=2
n=1
n=0
n=1
n=1
n=4
n=1
n=10
n=3
n=45
n=6
FDA cutoff
nmCRC
mCRC
nmCRC
mCRC
nmCRC
mCRC
MSS
MSI-L
MSI
TN
FP
TN
FP
TP
FN
0
25
50
75
100
Percentage

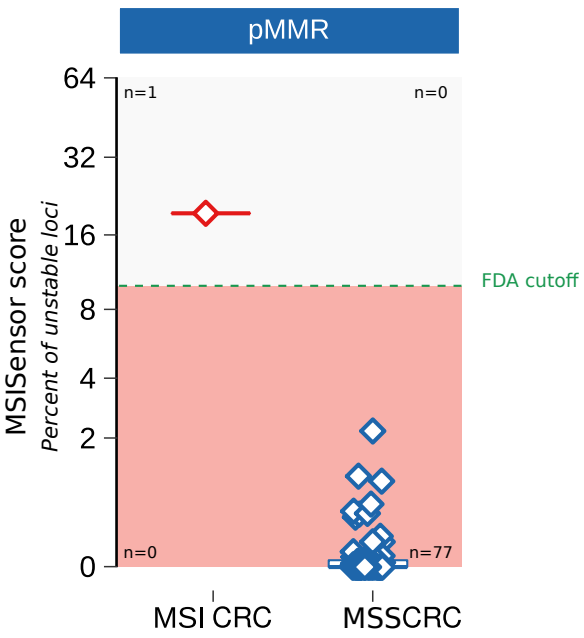

Status
MSI CRC
MSI-L CRC
MSS CRC
Tumor Stage
nmCRC
mCRC
MSISensor score
Percent of unstable loci
⩽10
>10

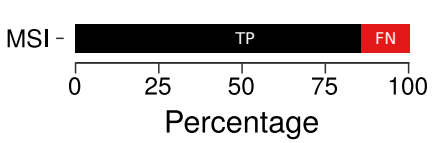
B

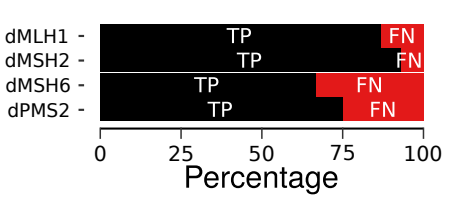

Retrospective (C2, n=113)
mCRC & nmCRC
MSI
MSISensor score
Percent of unstable loci
64
32
16
8
4
2
0
n=80
n=17
n=8
n=8
FDA cutoff
nmCRC
mCRC

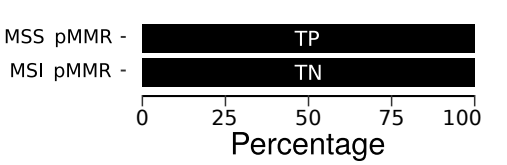
D


Prospective & Retrospective (C1+C2)
according to MMR deficiency
dMLH1
dMSH2
dMSH6
dPMS2
MSISensor score
Percent of unstable loci
64
32
16
8
4
2
0
n=79
n=27
n=2
n=3
n=12
n=2
n=1
n=1
FDA cutoff
MSI CRC
MSI CRC
MSI CRC
MSI CRC

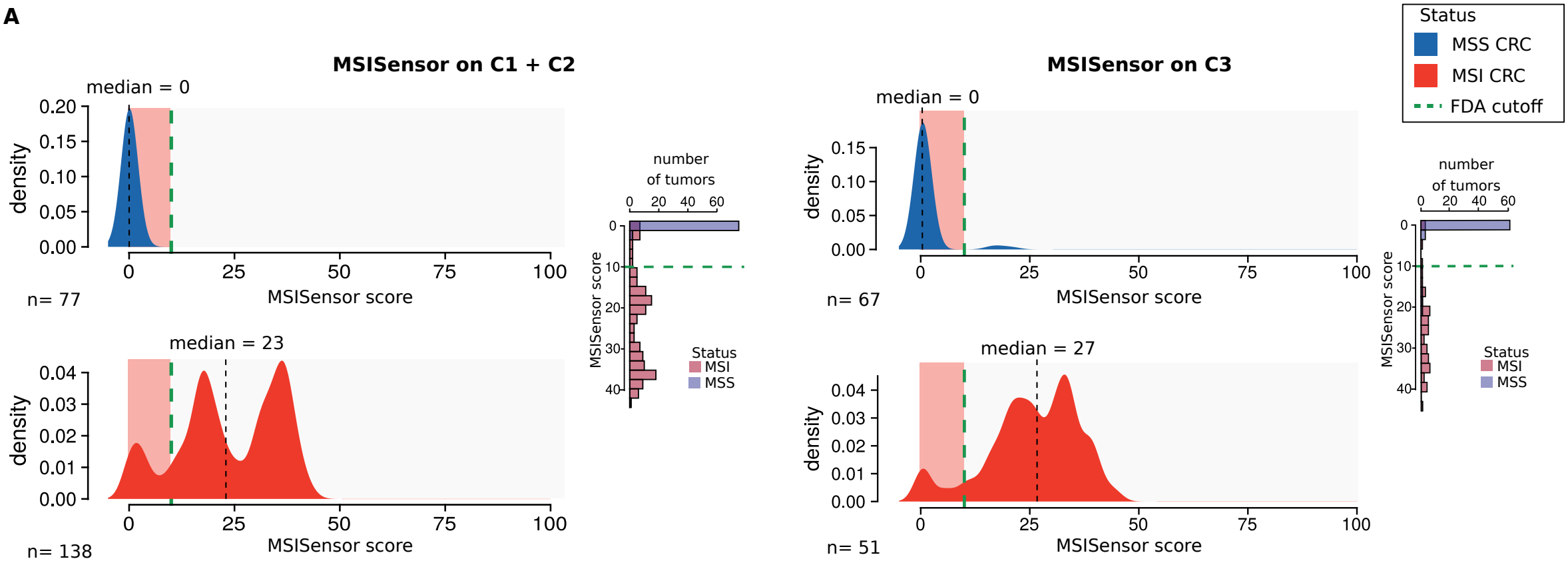

pMMR
MSISensor score
Percent of unstable loci
64
32
16
8
4
2
0
n=1
n=0
n=0
n=77
FDA cutoff
MSI CRC
MSSCRC


MSI
TP
FN
0
25
50
75
100
Percentage

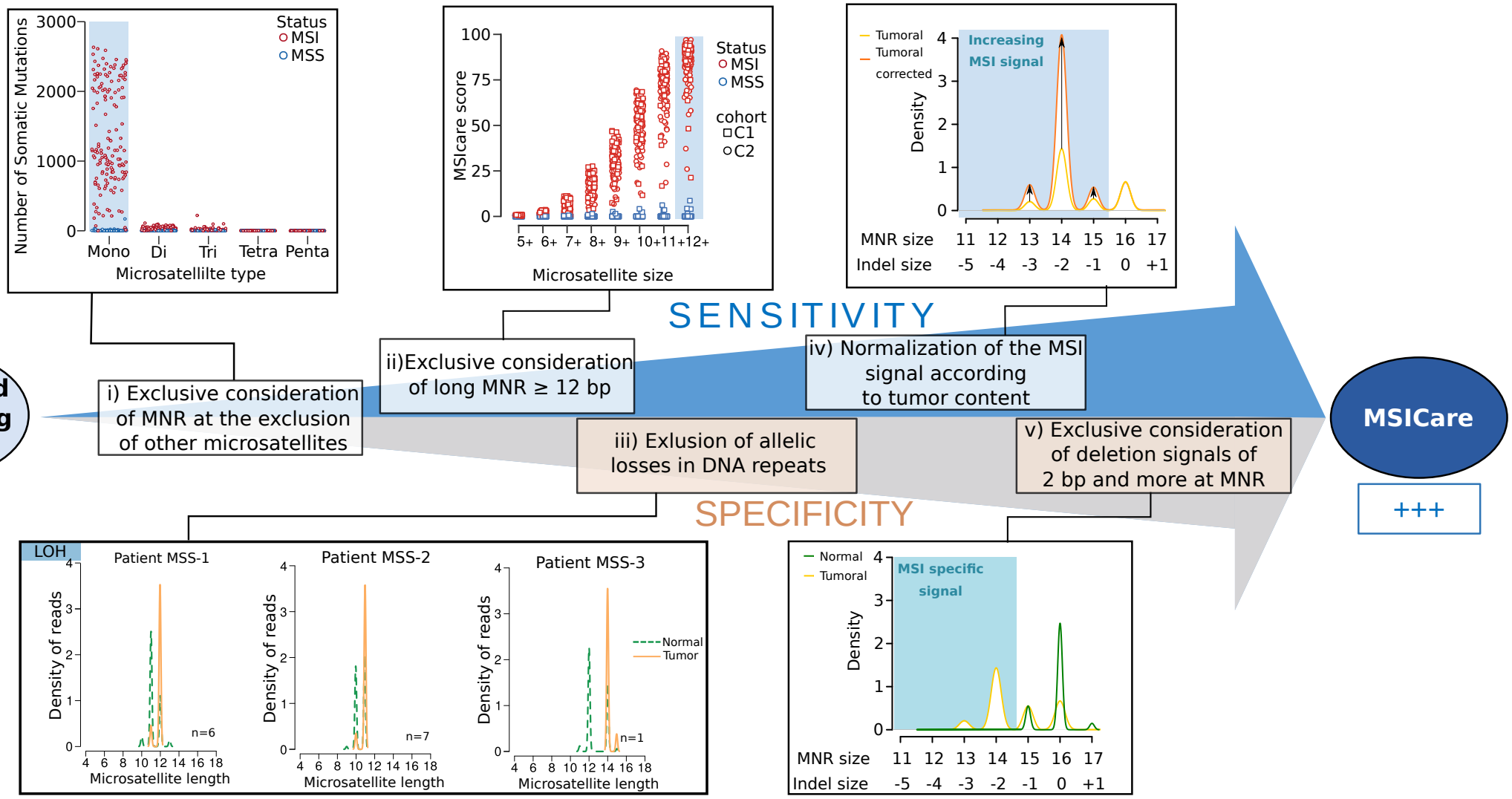

dMLH1
dMSH2
dMSH6
dPMS2
TP
FN
TP
FN
TP
FN
TP
FN
0
25
50
75
100
Percentage

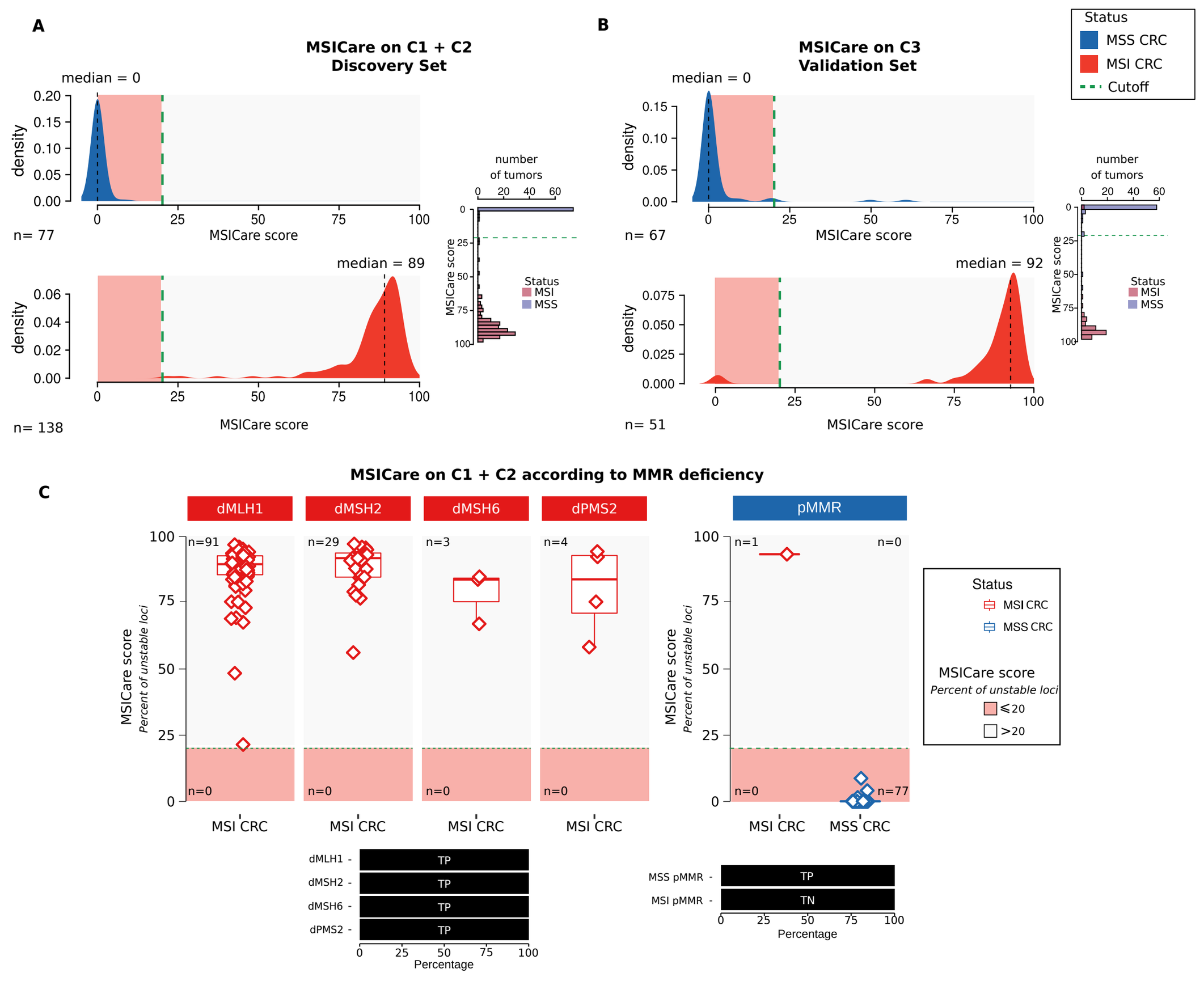

MSS pMMR
MSI pMMR
TP
TN
0
25
50
75
100
Percentage

A


MSISensor on C1 + C2
median = 0
density
MSISensor score
n= 77
median = 23
n= 138
number of tumors
MSISensor score
Status
MSI
MSS
MSISensor on C3
median = 0
n= 67
median = 27
n= 51
Status
MSS CRC
MSI CRC
FDA cutoff

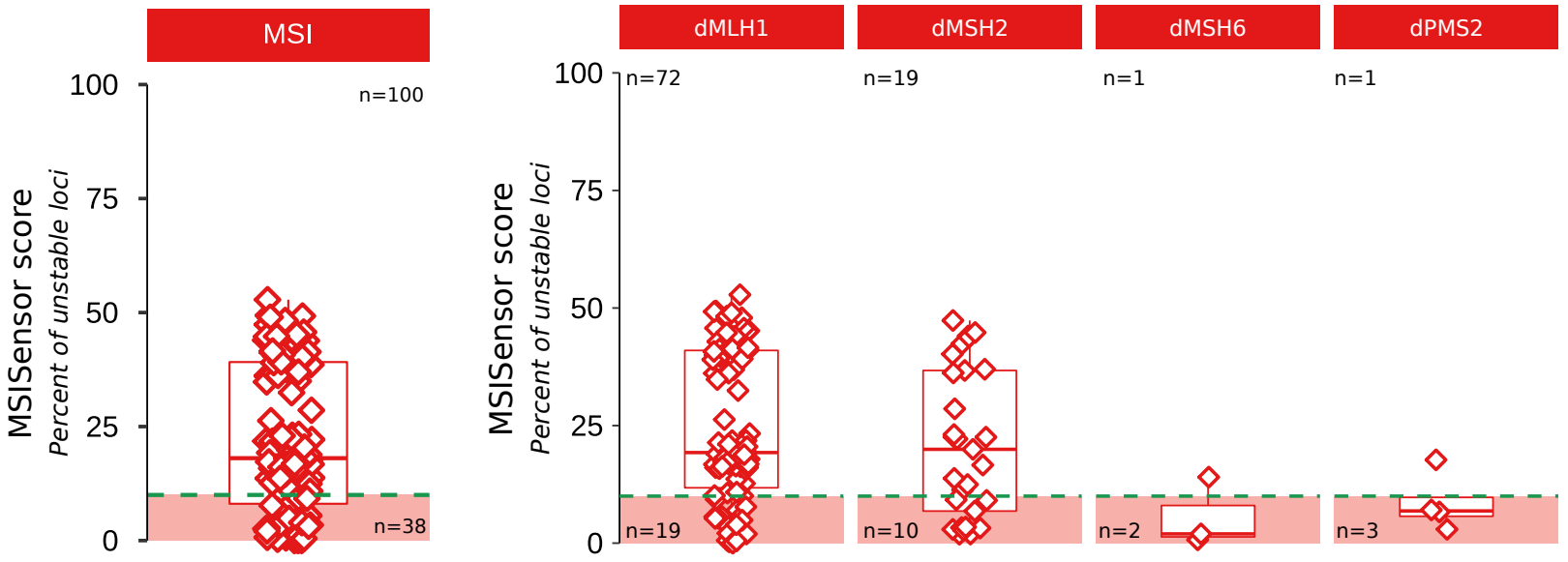
B

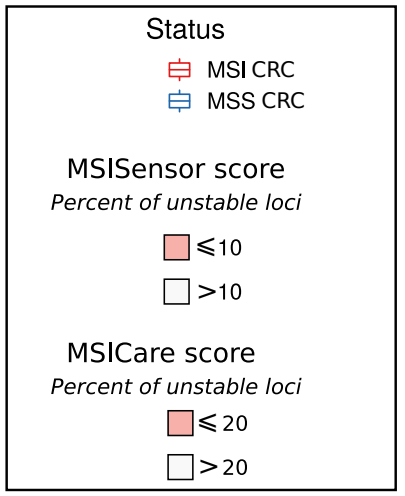

Number of Somatic Mutations
Mono
Di
Tri
Tetra
Penta
Microsatellilte type
MSIcare score
Microsatellite size
Tumoral
Tumoral corrected
Increasing MSI signal
Density
MNR size
Indel size
NGS-based MSI testing in CRC
-
i) Exclusive consideration of MNR at the exclusion of other microsatellites
ii)Exclusive consideration of long MNR ≥ 12 bp
iii) Exlusion of allelic losses in DNA repeats
iv) Normalization of the MSI signal according to tumor content
v) Exclusive consideration of deletion signals of 2 bp and more at MNR
SENSITIVITY
SPECIFICITY
MSICare
+++
LOH
Patient MSS-1
Patient MSS-2
Patient MSS-3
Density of reads
Microsatellite length
n=6
n=7
n=1
Normal
Tumor
MSI specific signal

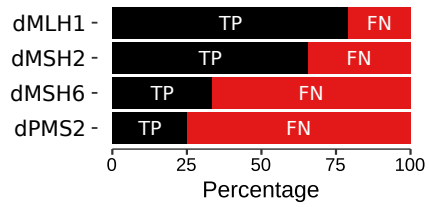

A
MSICare on C1 + C2
Discovery Set
median = 0
density
MSICare score
n= 77
median = 89
n= 138
number
of tumors
MSICare score
Status
MSI
MSS
B
MSICare on C3
Validation Set
median = 0
n= 67
median = 92
n= 51
Status
MSS CRC
MSI CRC
Cutoff
C
MSICare on C1 + C2 according to MMR deficiency
dMLH1
dMSH2
dMSH6
dPMS2
pMMR
n=91
n=29
n=3
n=4
n=1
n=0
n=77
MSICare score
Percent of unstable loci
MSI CRC
MSS CRC
MSICare score
Percent of unstable loci
⩽20
>20
TP
TN
Percentage

A

MSISensor in C1 and C2 with partial WES data restricted to MSK-IMPACT™

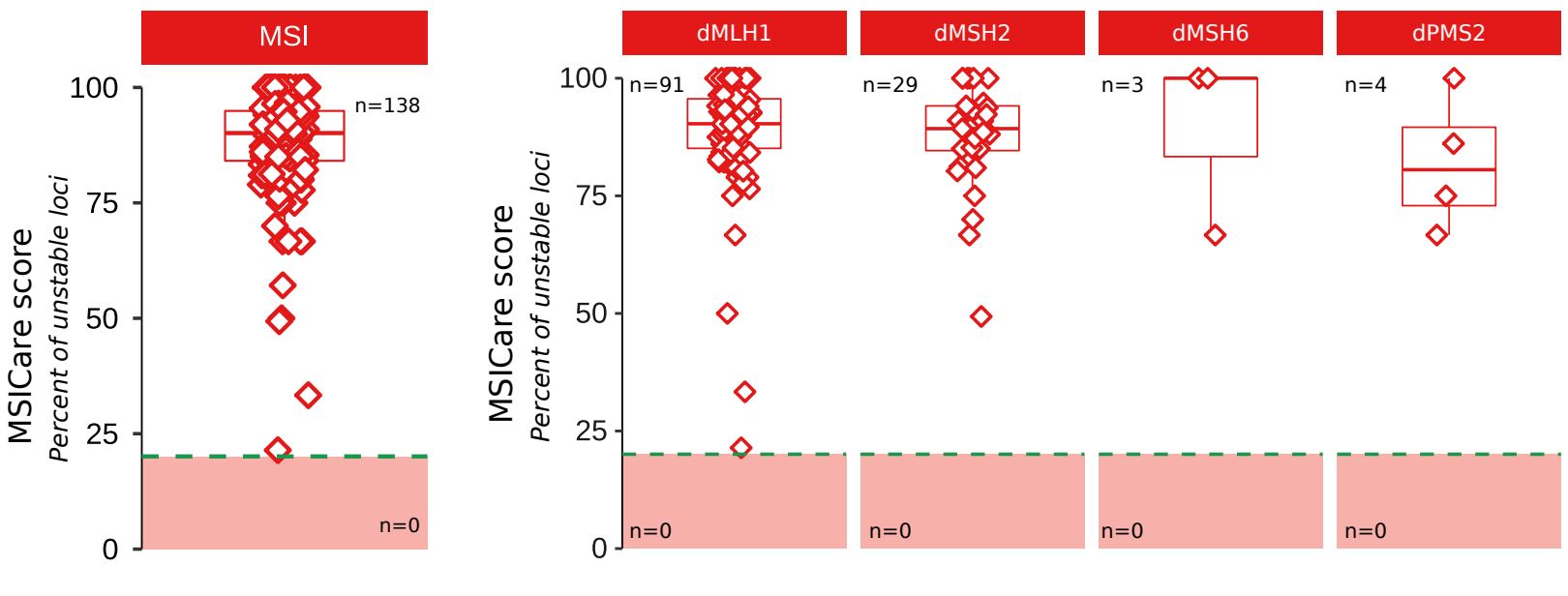


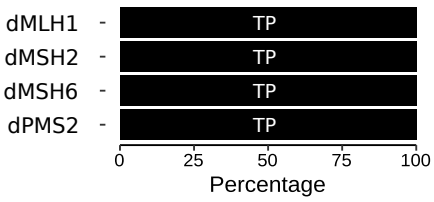




B

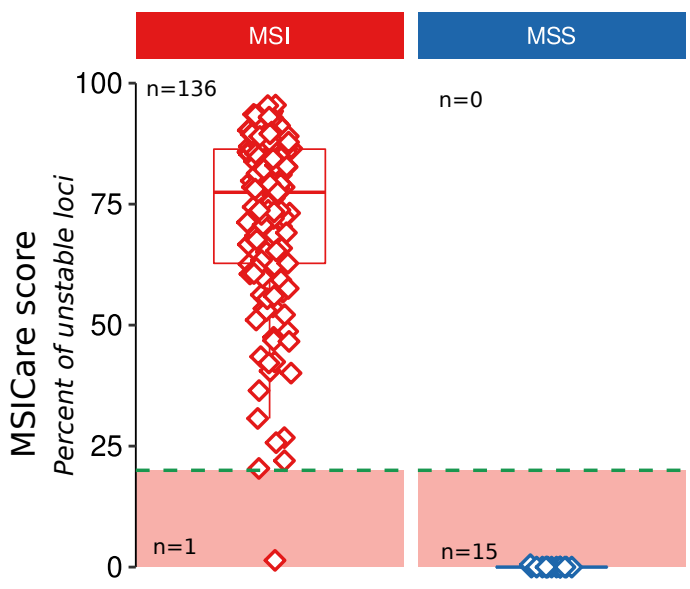
MSICare in C1 and C2 with partial WES data restricted to MSK-IMPACT™

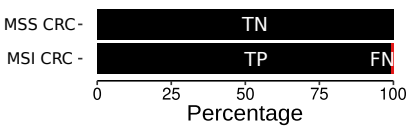


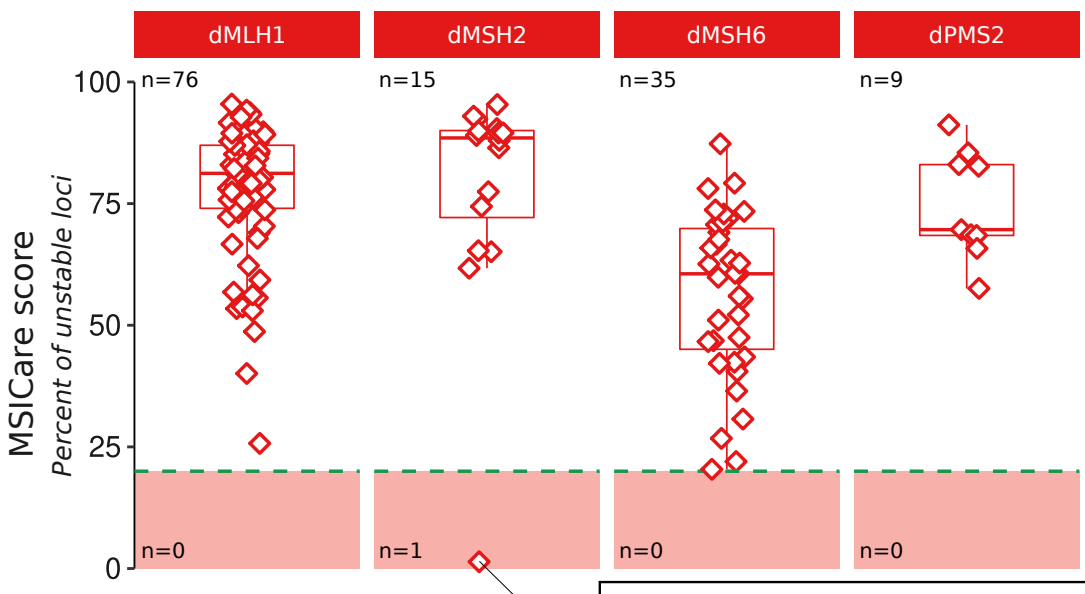


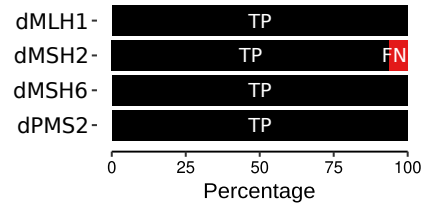
C

MSICare in C4 using targeted panel sequencing









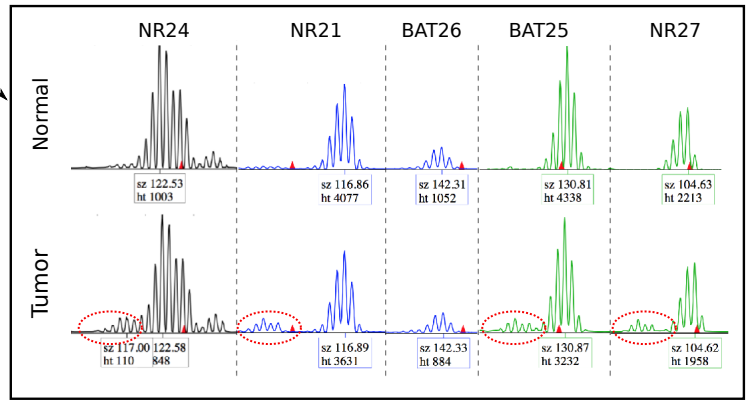


Table 1. Estimate of the sensitivity, specificity, negative and positive predictive values of microsatellite instability detection using MSIsensor.

	Sensitivity % (95 % CI)	Specificity % (95 % CI)	PPV % (95 % CI)	NPV % (95 % CI)
Method : MSISensor				
C1	84.0 (63.9 – 95.5)	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)	95.1 (87.8 – 98.6)
C2 ALL	85.8 (78.0 – 91.7)	NA	NA	NA
C2 nmCRC	90.9 (82.9 – 96.0)	NA	NA	NA
C2 mCRC	68.0 (46.5 – 85.1)	NA	NA	NA
C1+C2	85.5 (78.5 – 90.9)	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)	79.4 (70.0 – 86.9)
C3 CRC	90.2 (78.6 – 96.7)	95.5 (87.5 – 99.1)	93.9 (83.1 – 98.7)	92.8 (83.9 – 97.6)
C4	97.1 (92.7 – 99.2)	73.3 (44.9 – 92.2)	97.1 (92.7 – 99.2)	73.3 (44.9 – 92.2)

CI, confidence interval

PPV, Positive predictive value

NPV, Negative predictive value

nmCRC, non-metastatic colorectal cancer

mCRC, metastatic colorectal cancer

Table 2. Estimate of the sensitivity, specificity, negative and positive predictive values of microsatellite instability detection using MSIcare.

	Sensitivity % (95 % CI)	Specificity % (95 % CI)	PPV % (95 % CI)	NPV % (95 % CI)
Method : MSICare				
C1	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)
C2 ALL	100.0 (100.0 – 100.0)	NA	NA	NA
C2 nmCRC	100.0 (100.0 – 100.0)	NA	NA	NA
C2 mCRC	100.0 (100.0 – 100.0)	NA	NA	NA
C1+C2	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)
C3 CRC	96.1 (86.5 – 99.5)	97.0 (89.6 – 99.6)	96.1 (86.5 – 99.5)	97.0 (89.6 – 99.6)
C4	99.3 (96.0 – 100.0)	100.0 (100.0 – 100.0)	100.0 (100.0 – 100.0)	93.8 (69.8 – 99.8)

CI, confidence interval

PPV, Positive predictive value

NPV, Negative predictive value

nmCRC, non-metastatic colorectal cancer

mCRC, metastatic colorectal cancer