

**Manuscript version: Author's Accepted Manuscript**

The version presented in WRAP is the author's accepted manuscript and may differ from the published version or Version of Record.

**Persistent WRAP URL:**

<http://wrap.warwick.ac.uk/143562>

**How to cite:**

Please refer to published version for the most recent bibliographic citation information. If a published version is known of, the repository item page linked to above, will contain details on accessing it.

**Copyright and reuse:**

The Warwick Research Archive Portal (WRAP) makes this work by researchers of the University of Warwick available open access under the following conditions.

© 2020 Elsevier. Licensed under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International <http://creativecommons.org/licenses/by-nc-nd/4.0/>.



**Publisher's statement:**

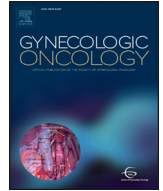
Please refer to the repository item page, publisher's statement section, for further information.

For more information, please contact the WRAP Team at: [wrap@warwick.ac.uk](mailto:wrap@warwick.ac.uk).



Contents lists available at ScienceDirect

Gynecologic Oncology

journal homepage: [www.elsevier.com/locate/gygno](http://www.elsevier.com/locate/gygno)

## Review Article

## Testing for lynch syndrome in people with endometrial cancer using immunohistochemistry and microsatellite instability-based testing strategies – A systematic review of test accuracy

Chris Stinton<sup>a</sup>, Hannah Fraser<sup>a</sup>, Lena Al-Khudairy<sup>a</sup>, Rachel Court<sup>a</sup>, Mary Jordan<sup>a</sup>, Dimitris Grammatopoulos<sup>a,b</sup>, Sian Taylor-Phillips<sup>a,\*</sup>

<sup>a</sup> Warwick Medical School, University of Warwick, Coventry CV4 7AL, United Kingdom

<sup>b</sup> Institute of Precision Diagnostics and Translational Medicine, University Hospital Coventry and Warwickshire (UHCW) NHS Trust, Coventry CV2 2DX, UK

## HIGHLIGHTS

- People with Lynch syndrome are at an increased risk of a range of cancers.
- This is the first systematic review of test accuracy of testing for Lynch syndrome in women with endometrial cancer.
- The sensitivity of the index tests were generally high, though most studies had much lower specificity.
- There was no evidence that test accuracy differed between testing strategies.

## ARTICLE INFO

## Article history:

Received 18 August 2020

Accepted 5 October 2020

Available online xxxx

## Keywords:

Systematic review

Lynch syndrome

Endometrial cancer

Immunohistochemistry

Microsatellite instability-based testing

Test accuracy

MLH1 promoter methylation testing

## ABSTRACT

**Background.** Lynch syndrome is an inherited genetic condition that is associated with an increased risk of cancer, including endometrial and colorectal cancer. We assessed the test accuracy of immunohistochemistry and microsatellite instability-based testing (with or without MLH1 promoter methylation testing) for Lynch syndrome in women with endometrial cancer.

**Methods.** We conducted a systematic review of literature published up to August 2019. We searched bibliographic databases, contacted experts and checked reference lists of relevant studies. Two reviewers conducted each stage of the review.

**Results.** Thirteen studies were identified that included approximately 3500 participants. None of the studies was at low risk of bias in all domains. Data could not be pooled due to the small number of heterogeneous studies. Sensitivity ranged from 60.7–100% for immunohistochemistry, 41.7–100% for microsatellite instability-based testing, and 90.5–100% for studies combining immunohistochemistry, microsatellite instability-based testing, and MLH1 promoter methylation testing. Specificity ranged from 60.9–83.3% (excluding 1 study with highly selective inclusion criteria) for immunohistochemistry, 69.2–89.9% for microsatellite instability-based testing, and 72.4–92.3% (excluding 1 study with highly selective inclusion criteria) for testing strategies that included immunohistochemistry, microsatellite instability-based testing, and MLH1 promoter methylation. We found no statistically significant differences in test accuracy estimates (sensitivity, specificity) in head-to-head studies of immunohistochemistry versus microsatellite instability-based testing. Reported test failures were rare.

**Conclusions.** Sensitivity of the index tests were generally high, though most studies had much lower specificity. We found no evidence that test accuracy differed between IHC and MSI based strategies. The evidence base is currently small and at high risk of bias.

© 2020 Published by Elsevier Inc.

**Abbreviations:** CI, Confidence interval; DNA, Deoxyribonucleic acid; FN, False negative; FP, False positive; IHC, Immunohistochemistry; IQR, Interquartile range; MHL1, MutL homologue 1; MLPA, Multiplex ligation-dependent probe; MSH2, MutS homologue 2; MSH6, MutS homologue 6; MSI, Microsatellite instability-based testing; MMR, Mismatch repair; NGS, Next-generation sequencing; NICE, National Institute for Health and Care Excellence; NPV, Negative predictive value; PMS2, Postmeiotic segregation increased 2; PPV, Positive predictive value; QUADAS-2, Quality Assessment Tool for Diagnostic Accuracy Studies 2; TN, True negative; TP, True positive.

\* Corresponding author.

E-mail address: [s.taylor-phillips@warwick.ac.uk](mailto:s.taylor-phillips@warwick.ac.uk) (S. Taylor-Phillips).

<https://doi.org/10.1016/j.ygyno.2020.10.003>

0090-8258/© 2020 Published by Elsevier Inc.

Please cite this article as: C. Stinton, H. Fraser, L. Al-Khudairy, et al., Testing for lynch syndrome in people with endometrial cancer using immunohistochemistry and microsat..., Gynecologic Oncology, <https://doi.org/10.1016/j.ygyno.2020.10.003>

## Contents

1.	Introduction . . . . .	0
2.	Methods . . . . .	0
2.1.	Search strategy . . . . .	0
2.2.	Eligibility criteria . . . . .	0
2.3.	Screening and data extraction . . . . .	0
2.4.	Quality appraisal . . . . .	0
2.5.	Data summary and synthesis . . . . .	0
3.	Results . . . . .	0
3.1.	Searching, sifting, and sorting. . . . .	0
3.2.	Quality appraisal . . . . .	0
3.2.1.	Complete test accuracy . . . . .	0
3.2.2.	Partial test accuracy studies . . . . .	0
3.3.	Characteristics of included studies . . . . .	0
3.4.	Description of screening and diagnostic tests . . . . .	0
3.5.	Accuracy of screening tests . . . . .	0
3.6.	Complete test accuracy studies . . . . .	0
3.6.1.	Head-to-head studies . . . . .	0
3.6.2.	Immunohistochemistry alone . . . . .	0
3.6.3.	Microsatellite instability-based testing alone. . . . .	0
3.6.4.	Immunohistochemistry and microsatellite instability-based testing, with MLH1 promoter methylation testing . . . . .	0
3.7.	Partial test accuracy studies . . . . .	0
3.7.1.	Immunohistochemistry alone . . . . .	0
3.7.2.	Immunohistochemistry with MLH1 promoter methylation testing . . . . .	0
3.7.3.	Microsatellite instability-based testing alone. . . . .	0
3.7.4.	Immunohistochemistry and microsatellite instability-based testing, with MLH1 promoter methylation testing . . . . .	0
4.	Discussion. . . . .	0
5.	Conclusions . . . . .	0
	Definition of test accuracy terms . . . . .	0
	Declarations . . . . .	0
	Consent for publication . . . . .	0
	Availability of data and materials . . . . .	0
	Funding . . . . .	0
	Author contributions . . . . .	0
	Ethics statement . . . . .	0
	Declaration of Competing Interest . . . . .	0
	. . . . .	0
	Appendix A. Supplementary Data . . . . .	0
	References . . . . .	0

## 1. Introduction

Endometrial cancer is the most common gynaecological cancer in the Western world [1]. The incidence of endometrial cancer generally increases with age, reaching a peak of 97.3 per 100,000 population between the ages of 75 and 79 years [2]. A recent estimate suggests that people with endometrial cancer have a 1-year survival rate of 89.6% and a 5-year survival rate of 75.7% [3]. Risk factors for the development of endometrial cancer include obesity, nulliparity, early age at menarche, use of hormone-replacement therapy, and gene mutations (e.g. Lynch syndrome) [4–6].

Lynch syndrome is an autosomal dominant cancer-predisposition syndrome that is most frequently caused by mutations to any of four mismatch repair (MMR) genes: MutL homologue 1 (MLH1), MutS homologue 2 (MSH2), MutS homologue 6 (MSH6), or postmeiotic segregation increased 2 (PMS2) [7]. MMR genes encode proteins that are involved in recognising and repairing errors that occur in DNA during cell division. Mutations in MMR genes prevent DNA errors from being corrected, which can result in uncontrolled cell growth and the development of cancer. The most common cancers among people with Lynch syndrome are colorectal and endometrial cancers [8]. Detection of Lynch syndrome might lead to reductions in the risk of developing cancer for both the individual and their family members (through

surveillance and risk-reducing strategies such as chemoprevention) and earlier treatment of cancers. Two main tests are used to identify people who are at higher risk of Lynch syndrome: immunohistochemistry (IHC) and microsatellite instability (MSI)-based testing. Additional testing for MLH1 promoter hypermethylation can assist in the identification of sporadic cancers. Universal testing for Lynch syndrome in people with endometrial cancer has been recommended by professional organisations in the Europe and the US. [9,10] The present review was commissioned by NICE in the UK to determine whether to conduct testing there.

A recent systematic review of IHC and MSI for Lynch syndrome in colorectal cancer has suggested sensitivity and specificity are variable for each of the tests [11]. No such systematic assessment has been conducted in endometrial cancer. The aim of this systematic review is to assess the test accuracy and test failure rates of IHC and MSI-based strategies for detecting Lynch syndrome in people who have a diagnosis of endometrial cancer.

## 2. Methods

The review is registered on the PROSPERO database (registration number CRD42019147185) and the protocol is available from the NICE website ([www.nice.org.uk/guidance/indevelopment/gid-dg10033](http://www.nice.org.uk/guidance/indevelopment/gid-dg10033)).

## 2.1. Search strategy

The search strategy comprised the following main elements:

- 1) Searching of electronic bibliographic databases,
- 2) Contacting experts in the field, and
- 3) Scrutiny of references of included studies and relevant systematic reviews.

Searches were developed iteratively and conducted in the following databases, from inception: MEDLINE (Ovid); Embase (Ovid); Cochrane Database of Systematic Reviews and *Cochrane Central Register of Controlled Trial* (Wiley); Database of Abstracts of Reviews of Effects and Health Technology Assessment database (CRD), Science Citation Index and Conference Proceedings (Web of Science), and PROSPERO International Prospective Register of Systematic Reviews (CRD). Search terms related to endometrial cancer and Lynch syndrome. Full details of the search strategies are provided in Supplement 1. Searches were undertaken in August 2019.

## 2.2. Eligibility criteria

We included English language journal articles that investigated testing for Lynch syndrome by immunohistochemistry or microsatellite instability-based testing (with or without MLH1 promoter hypermethylation testing) in women who have endometrial cancer. The primary reference standard was genetic verifications of constitutional mutations in the MMR genes by sequencing/next-generation sequencing (with or without multiplex ligation-dependent probe amplification). Other acceptable reference standards were array-based comparative genomic hybridization, and long-range polymerase chain reaction. Outcomes were any reported test accuracy measure, or test failures. We excluded

studies of women with pre-cancerous conditions of the uterus (i.e. atypical endometrial hyperplasia), articles not available in English, non-human studies, letters, reviews, editorials, grey literature, conference abstracts, and communications. We also excluded studies without extractable numerical data or with insufficient information for quality appraisal, studies where more than 10% of the sample do not meet our inclusion criteria, and studies in which fewer than 95% of women who had the index test (or who were index test positive for studies with incomplete test accuracy results) had germline testing (to minimise sampling biases).

## 2.3. Screening and data extraction

Two reviewers independently screened titles, abstracts, full text papers, and extracted data. At each stage of the review, disagreements were resolved through discussion between the reviewers, with the involvement of a third reviewer if required.

## 2.4. Quality appraisal

Quality appraisal was assessed independently by two reviewers using the Quality Assessment Tool for Diagnostic Accuracy Studies (QUADAS-2) [12]. We tailored the tool to the research question. Tailoring comprised (1) adding a domain to assess whether quality assurance measures were in place, (2) defining inappropriate participant exclusion criteria, (3) identifying appropriate reference standards, and (4) adding a domain on the role of the sponsors. We also produced guidance notes. Disagreements were resolved through discussion between the two reviewers, with the involvement of a third reviewer if required.

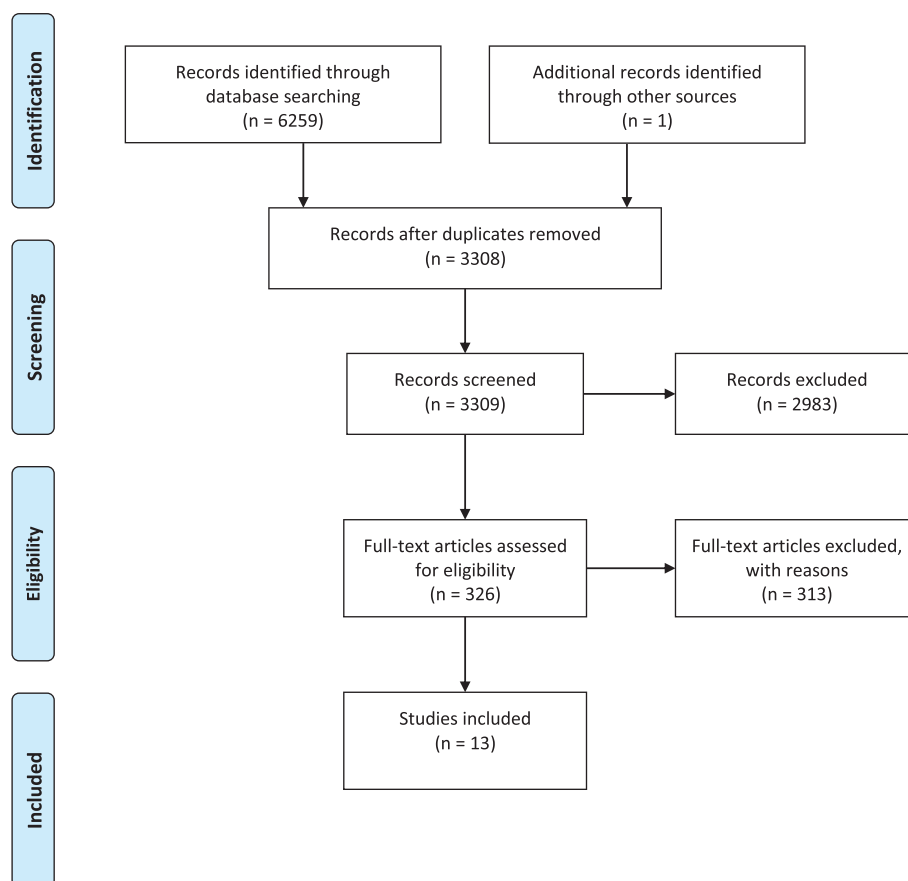


Fig. 1. PRISMA flow diagram.

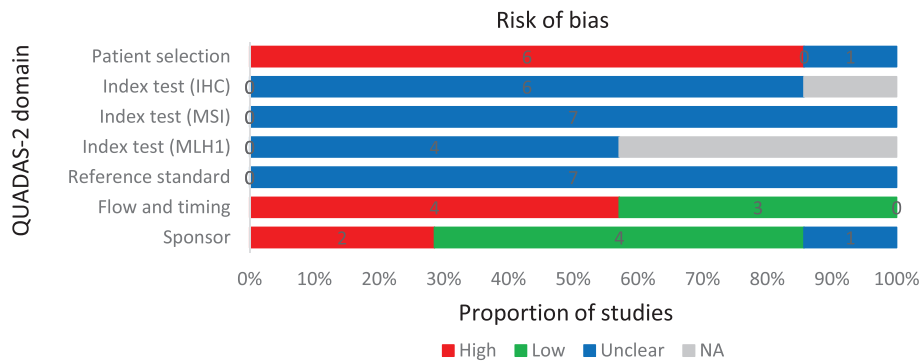


Fig. 2. Risk of bias in complete test accuracy studies.

A copy of the QUADAS-2 tool and guidance notes are provided in supplements 2 and 3, respectively.

2.5. Data summary and synthesis

In an ideal test accuracy study, an entire sample of participants receives both the index test and the reference standard. This facilitates direct comparisons of the agreement between the two tests, and minimises biases that might occur as result of participant characteristics differing at each stage of the process. For reasons of costs, practicality, and ethics, in some test accuracy studies only a subset of participants receive both tests, i.e. only those who are index test positive will be offered the reference standard; participants who are index test negative receive no further testing. This approach accurately reflects how tests are used in clinical practice, but can lead to inaccurate estimates of test accuracy as the true diagnostic status of participants who are index test negative is not known. This ‘partial verification bias’ has been found to overestimate sensitivity and underestimate specificity [13]. In this paper, we divide results in complete test accuracy studies (in which participants have received both the index test and reference standard) and partial test accuracy studies (in which only participants who are index test negative receive the reference standard). For complete test accuracy studies, we extracted data on true positives, false positives, false negatives, and true negatives and calculated sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). For partial test accuracy studies (in which sensitivity, specificity, and NPV could not be calculated), we extracted true positives and false positives, and calculated PPV. Confidence intervals were calculated using Wilson’s continuity correction [14].

Meta-analysis of test accuracy was not possible due to the small number of heterogeneous studies. We provide a narrative summary of results, reporting test accuracy estimates and presenting forest plots. Summary statistics and forest plots were generated using Stata 16 [15].

3. Results

3.1. Searching, sifting, and sorting

Fig. 1 provides full details of the flow of studies through the review. We identified 3308 unique records through electronic databases. One additional study was provided after contact with experts [16]. After examination of titles and abstracts, 326 papers were retained for full text assessment. Thirteen papers met the review’s inclusion criteria and were included in the review (7 complete test accuracy, 6 partial test accuracy). Supplement 4 contains a list of excluded studies (with reasons for exclusions).

3.2. Quality appraisal

Figs. 2 and 3 provide summaries of risks of bias of the included studies. Quality appraisal for each individual study is shown in Supplement 5.

3.2.1. Complete test accuracy

Risk of bias was high in two or more domains for six studies (85.7%) [17–22]. The remaining study was at unclear risk of bias in five domains, and low risk of bias in two domains [23]. No study was at low risk of bias in all domains. In the patient selection domain, there was a high risk of bias in 85.7% of studies (6/7 studies) [17–22]. In the index test domain,

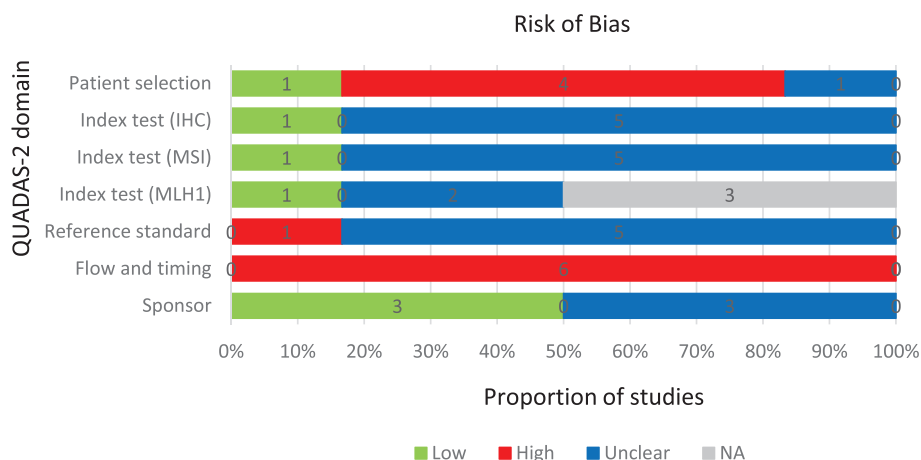


Fig. 3. Risk of bias in partial test accuracy studies.

**Table 1**  
Characteristics of included studies.

Study reference	Country	Study design	Study setting	Time period	Sample size included	Selected/unselected sample	Age Mean Median (range)	Ethnicity	Previous /concurrent cancers	Relatives	Index test(s)	Reference standard tests(s)
Berends 2003	Netherlands	Retrospective and prospective cohort	Cancer registry	Before 1989–2000	58	Selected	Median 45 (27–49) years	NR	13/38 (22.4%)	22/58 (37.9%) cancer diagnosis in 1st degree relatives	MSI and IHC	DGGE and sequencing
Chao 2019	China	Prospective cohort	Hospital	Dec 2017–Aug 2018	111	Selected	Mean 55.7 years Median 55 (31–82) years	NR	0 – excluded	14/111 (12.6%) Amsterdam II criteria, 2 met Bethesda criteria	IHC, MSI and MLH1 promoter hypermethylation testing	NGS and Sanger sequencing
Goodfellow 2015	USA	Propsective cohort	Hospital	2003–2007	1043	Selected after 2007	Mean 62 (25–100) years	White, <i>n</i> = 848 (90.4%) African American, <i>n</i> = 55 (5.9%) Asian, <i>n</i> = 17 (1.8%) Other, <i>n</i> = 7 (0.7%) Unknown/not specified, <i>n</i> = 11 (1.2%)	NR	938/1043 (90%) had Lynch associated cancers	MSI, IHC and MLH1 promoter hypermethylation testing	NGS
Latham 2019	USA	Retrospective Cohort	Hospital	Jan 2014–Jun 2017	525	Unclear	Median 55–60 years across all MSI groups	NR for whole sample	NR	NR	MSI and IHC	NGS
Lu 2007	USA	Prospective cohort	Gynaecologic oncology clinics	Jan 2000 end date NR	100	Selected	Mean 41.6 Median 43 (24–49) years	NR	12/100 (12%) 2 Colon 9 synchronous ovarian 1 brain	21/100 (21%) LS related cancer in at least 1 first degree relative	MSI, IHC and MLH1 promoter hypermethylation testing	Sequencing and unclear further testing for large deletions
Mercado 2012	USA	Retrospective cohort study	Hospitals	NR	129	Selected	Median 63 (38–89) years	94 (73%) Caucasian 1 (1%) Hispanic 1 (1%) Asian 2 (2%) other	34 (27%) CRC 6 (5%) adenoma 33 (26%) other Lynch 37 (29%) Multiple LS	115/129 (89%) CRC 48/129 (37%) EC 67 (52%) Other LS cancer	MSI, IHC	DHPLC and sequencing
Ollikainen 2005	Finland	Cohort (retrospective and prospective)	Hospital	1986–1997	23	Selected	Mean 62 years Median 61 (32–81) years	NR	2/23 (9%) breast cancer	23/23 (100%) family history of EC	MSI, IHC and MLH1 promoter hypermethylation testing	Sequencing and MLPA
Ring 2016	USA	Prospective cohort	Hospital	NR	381	Unselected adult only	Mean 61 years at diagnosis	Caucasian <i>n</i> = 265 (70%) African–American <i>n</i> = 34 (9%) Hispanic <i>n</i> = 66 (17%) Asian <i>n</i> = 14 (4%) Native American <i>n</i> = 2 (1%)	NR	NR for whole sample	MSI, IHC and MLH1 promoter hypermethylation testing	NGS and MLPA
Ryan 2020	UK	Prospective cohort	Hospital	2 years	500	Unselected	Median 65 years	White <i>n</i> = 405 (81%) Black <i>n</i> = 20 (4%) Asian <i>n</i> = 55 (11%) Chinese <i>n</i> = 10 (2%)	3/500 (0.6%)	NR for whole sample	IHC, MSI and MLH1 promoter hypermethylation testing	NGS and MLPA

(continued on next page)

Table 1 (continued)

Study reference	Country	Study design	Study setting	Time period	Sample size included	Selected/unselected sample	Age Mean Median (range)	Ethnicity	Previous /concurrent cancers	Relatives	Index test(s)	Reference standard tests(s)
Rubio 2016	Spain	Retrospective and prospective cohort	Hospital	3 years NR	103	Selected	NR	Other n = 10 (2%) NR	Colon, n = 20 (19.4%) Ovary, n = 14 (13.6%) Skin, n = 4 (3.9%)	64/99 (65%) available histories	MSI and IHC	CSGE sequencing, MLPA
Salvador 2019	USA	Retrospective cohort study	Laboratory/hospital	2016–2018	237	Selected	NR for EC patients alone	NR for EC patients alone	NR for EC sample alone	NR for EC sample alone	MSI, IHC and MLH1 promoter hypermethylation testing IHC	NGS and MLPA Sequencing, NGS and MLPA
Tian 2019	China	Propsective cohort	Cancer centre	Jan 2014–Jul 2017	198	Selected	NR in whole sample	Chinese	44/196 (22.4%) multiple primary tumour 20 CRC 6 ovarian	47/196 (24%) LS related tumour in a first degree relative		Sequencing promoter methylation testing
Yoon 2008	Korea	Prospective cohort	Hospital	Jan 1996–Dec 2004	113	Selected	NR	NR	NR	NGS and MLPA	MSI, IHC and MLH1 promoter methylation testing	Sequencing

CSGE = conformation sensitive gel electrophoresis, DGGE = denaturing gradient gel electrophoresis, DHPIC = Denaturing high performance liquid chromatography, EC = endometrial cancer, IHC = immunohistochemistry, MLPA = multiplex ligation-dependent probe amplification, MSI = microsatellite instability, NGS = next-generation sequencing, NR = not reported.



all studies were rated as having unclear risk of bias for IHC [17–23]. In the reference standard domain, risk of bias was rated as unclear in all seven studies [17–23]. The flow of patients through the study was rated at high risk of bias in 57% of studies (4/7 studies) [17,18,20,21]. Three of the studies did not include all patients in their analysis [18,20,21], and two studies did not give all patients the same reference standard as sequencing was only conducted after positive testing with conformation sensitive gel electrophoresis or denaturing gradient gel electrophoresis [18,20]. The remaining three studies had a low risk of bias [19,22,23]. Risk of bias due to the role of the sponsor was high in two studies in which multiple authors were employed by companies who conducting testing for Lynch syndrome and funded the studies [19,22].

3.2.2. Partial test accuracy studies

Risk of bias was high in two or more domains for 67% of the studies (4/6 studies) [24–27]. No study was at low risk of bias in all domains. The areas where there were a high risk of bias were patient selection (4/6 studies, 67%) [24–27], flow and timing (all studies) [16,24–28], and the reference standard (1/5 studies, 20%) [27]. There were high applicability concerns in the patient selection domain in three studies [24,25,27], with these studies narrowing their inclusion criteria by age and personal/familial cancer history. No study had high applicability concerns in the index test and reference standard domains.

3.3. Characteristics of included studies

Characteristics of the included studies are provided in Table 1. Of the 13 studies, seven studies provided complete test accuracy data [17–23], and six studies provided partial test accuracy data [16,24–28]. Across the 13 studies, there were approximately 3500 participants, ranging from 23 [25] to 1043 [26] patients per study. Four studies took place in European countries [16,17,20,25], three in Asian countries [18,21,24], and six in the USA [19,22,23,26–28]. One study was conducted in an unselected population [16], and 12 studies in selected populations (i.e. participants were limited to only those with particular characteristics, e.g. under 50 years only, without a person/family history of cancer) [17–28]. All studies within this review had a cohort design, 53.8% (6/13) of studies were prospective cohorts [16,18,19,21,23,24,26], 23.1% (3/13) were retrospective [22,27,28], and 23.1% (3/13) had both prospective and retrospective elements [17,20,25].

3.4. Description of screening and diagnostic tests

Ninety-two per cent (12/13) of studies included both IHC and MSI tests. One study used IHC only [21]. Sixty-two per cent (8/13) of studies also included MLH1 promoter methylation testing [16,18,19,22–26].

In studies of IHC, seven studies tested all four IHC proteins (MLH1, MSH2, MSH6 and PMS2) [16,18–22,27], and three studies tested three proteins (MLH1, MSH2, MSH6) [23–25]. IHC cut-offs were not reported

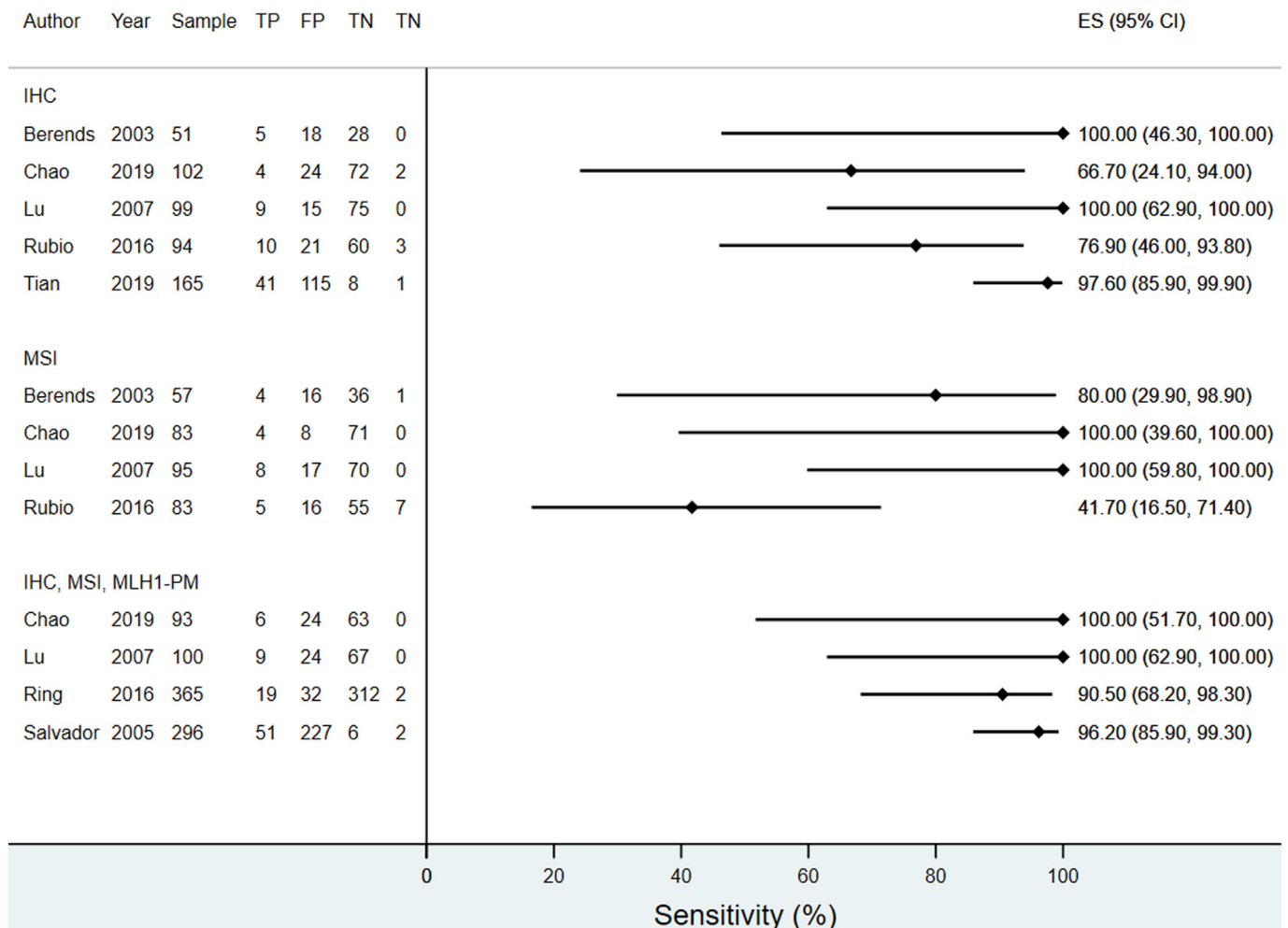


Fig. 4. Forest plots of complete test accuracy studies.



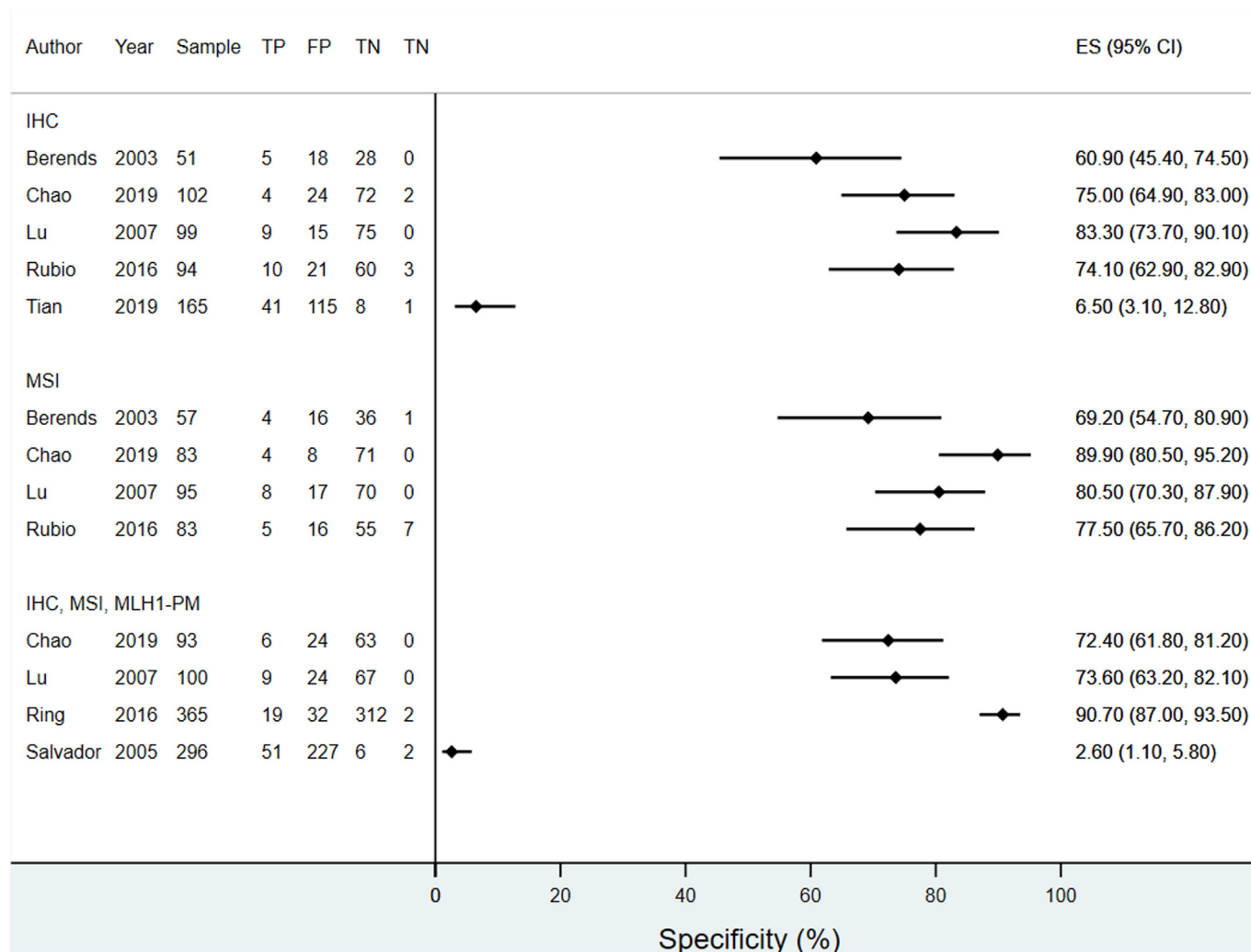


Fig. 4 (continued).

in two studies [20,21]. In studies of MSI, three different panels of markers were used: BAT25, BAT26, D2S123, D5S346, and D17S250 [17,24–26]; BAT25, BAT26, BAT40, D2S123, D5S346, and D17S250 [20,23]; and BAT-25, BAT-26, NR-21, NR-24, NR-27, and MONO-27 [16,18]. The panel of markers was not reported in four studies [19,22,27,28]. The circumstances under which MLH1-PM testing was conducted varied between studies. MLH1 promoter methylation testing was conducted if tumours were categorised as MSI-H or had IHC loss (MLH1 or MLH1/PMS2) in three studies [16,22,23], if tumours had IHC MLH1 loss only in three studies [18,24,25], and in all participants in one study [26]. In the remaining paper, the circumstances under which MLH1-PM was conducted was not reported [19].

Cut-offs for index tests varied across studies. For IHC, index positive was defined as complete absence of protein expression [19], any loss of expression [16–18,23,27], or no evidence of protein expression [24]. IHC cut-offs were not reported for the remaining five studies [20–22,24–28]. For MSI, index test positive was defined as MSI-H ( $\geq 2$  instable markers) in nine studies [16–18,22,23], and an MSI sensor score  $\geq 10$  in one study [28]. For the remaining two studies, the cut-off was described as MSI-H but the number of markers indicating instability was not reported [19,20].

A range of reference standards were employed: sequencing [24,27], next-generation sequencing (NGS) [26,28], sequencing and multiplex ligation-dependent probe amplification (MLPA) [16,17,19,20,22,25], sequencing and NGS [18], sequencing, NGS, and MLPA [21], and sequencing with an unspecified method to detect large deletions [23].

Three of the included studies used an additional reference standard test prior to sequencing: conformation sensitive gel electrophoresis [20], denaturing gel electrophoresis [17], and denaturing high performance liquid chromatography [27].

### 3.5. Accuracy of screening tests

The methods, thresholds to determine positivity of index tests, and the diagnostic tests varied between studies. Results were considered positive when they exceeded the threshold as set in the individual study. Forest plots of test accuracy metrics are given in Fig. 4 (complete test accuracy) and Fig. 5 (partial test accuracy). Test failures are reported in Supplement 6.

### 3.6. Complete test accuracy studies

#### 3.6.1. Head-to-head studies

Four studies provided data that could be used to explore head-to-head comparisons between IHC and MSI testing [17,18,20,23]. However, the number of tumours assessed by IHC and MSI in each studies were not identical. More IHC than MSI results were available in three studies (102 vs. 83 [18], 99 vs. 95 [23], and 94 vs. 83 [20]), and more MSI than IHC results were available in one study (57 vs. 51 [17]).

For IHC there was combined total of 28 true positives, 78 false positives, 235 true negatives, and 5 false negatives across the studies. For MSI there

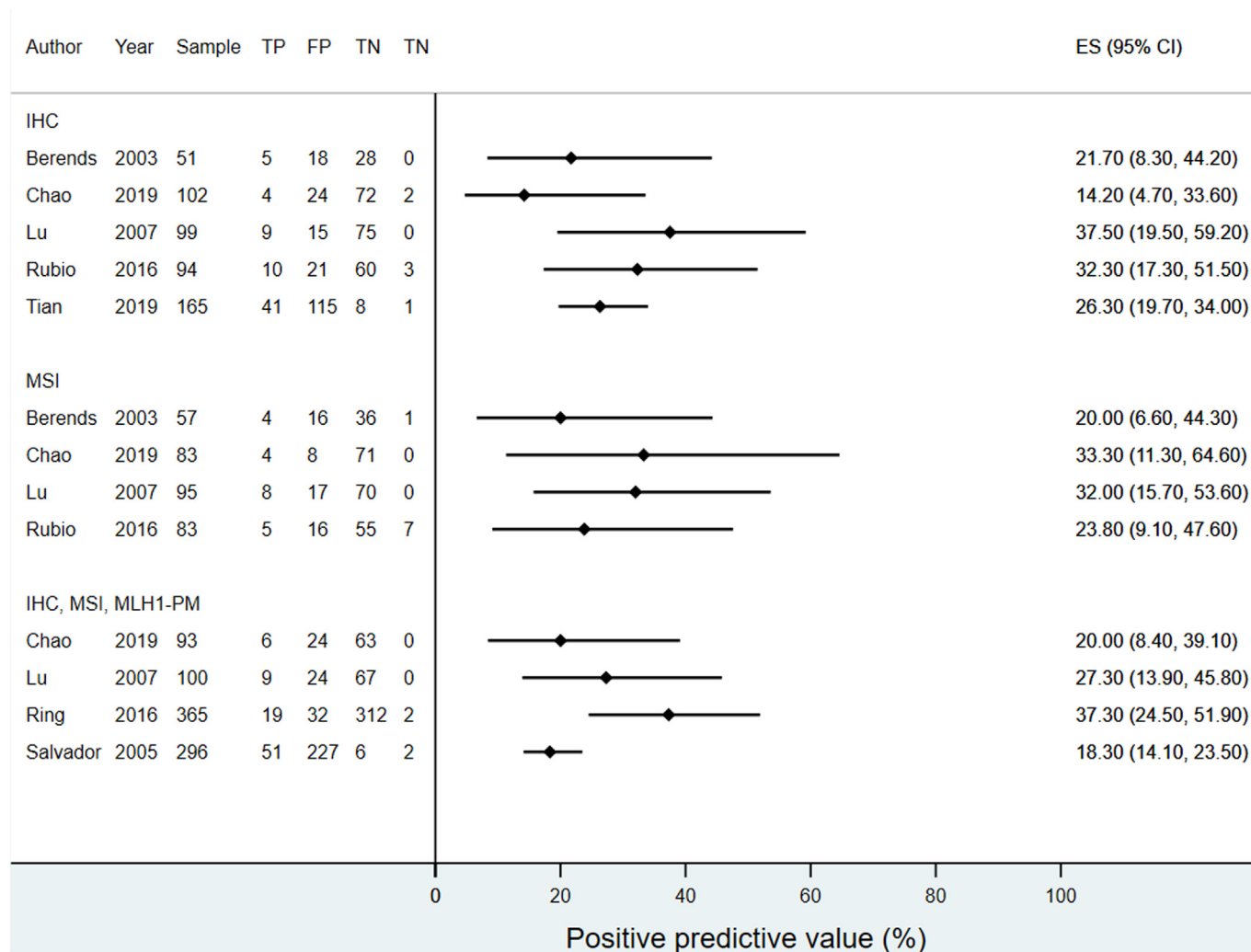


Fig. 4 (continued).

was a combined total 21 true positives, 57 false positives, 232 true negatives, and 8 false negatives in the studies. Median test accuracy estimates were similar between the two tests. For sensitivity, estimates ranged were 66.7–100% for IHC and 41.7–100% for MSI. For specificity, estimates were 60.9–83.3% for IHC and 69.2–89.9% for MSI. For PPV, estimates were 14.2–37.5% for IHC and 20–33.3% for MSI. For NPV, estimates were 95.2–100% for IHC and 88.8–100% for MSI. Within studies, confidence intervals of IHC and MSI on each of the four test accuracy estimates overlapped, suggesting no statistically significant differences between the tests.

### 3.6.2. Immunohistochemistry alone

Data on IHC testing alone were available from five studies [17,18,20,21,23]. There were 69 true positives, 193 false positive, 243 true negatives, and 6 false negatives. Test accuracy estimates ranged from 66.7–100% for sensitivity, 6.5–83.3% for specificity, 14.2–37.5% for PPV, and 88.9–100% for NPV. Excluding the study with highly selective inclusion criteria [21] narrowed the specificity range to 60.9–83.3%. Test failures were reported for 0–1% of tumours for immunohistochemistry (1 out of 522 tumours).

### 3.6.3. Microsatellite instability-based testing alone

Data on MSI testing alone using two or more markers as a cut-off were available from four studies [17,18,20,23]. There were 21 true positives, 57 false positive, 232 true negatives, and 8 false negatives. Test accuracy estimates ranged from 41.7–100% for sensitivity, 69.2–89.9% for specificity, 20–33.3% for PPV, and 88.8–100% for NPV. No test failures were reported.

Data on MSI testing using one or more markers as a cut-off were available from one study [20]. There were 5 true positives, 17 false positives, 54 true negatives, and 7 false negatives. Sensitivity was 41.7% (95% CI 16.5–71.4%), specificity was 76.1% (95% CI 64.2–85.1%), PPV was 22.7% (95% CI 8.7–45.8%), and NPV was 88.5% (95% CI 77.2–94.9%). Using a cut-off of one or more stable marker (rather than two or more) changed the status of one index test result from true negative to false positive. No test failures were reported.

### 3.6.4. Immunohistochemistry and microsatellite instability-based testing, with MLH1 promoter methylation testing

Four studies provided test accuracy data for immunohistochemistry and microsatellite instability-based testing [18,19,22,23]. There were 85 true positives, 290 false positives, 465 true negatives, and 4 false negatives. Test accuracy estimates ranged from 90.5–100% for sensitivity, 2.6–92.3% for specificity, 18.3–56.3% for PPV, and 75–100% for NPV. Excluding the study with highly selective inclusion criteria [22] narrowed the specificity range to 72.4–92.3%. Test failures were reported in one study: 0.2% (1/567 tumours) for IHC [23]. No test failures were reported for MSI or MLH1 promoter methylation testing.

### 3.7. Partial test accuracy studies

#### 3.7.1. Immunohistochemistry alone

Two studies provided test accuracy data for IHC alone [16,27]. One study assessed PPV for all proteins combined [16], and one study for

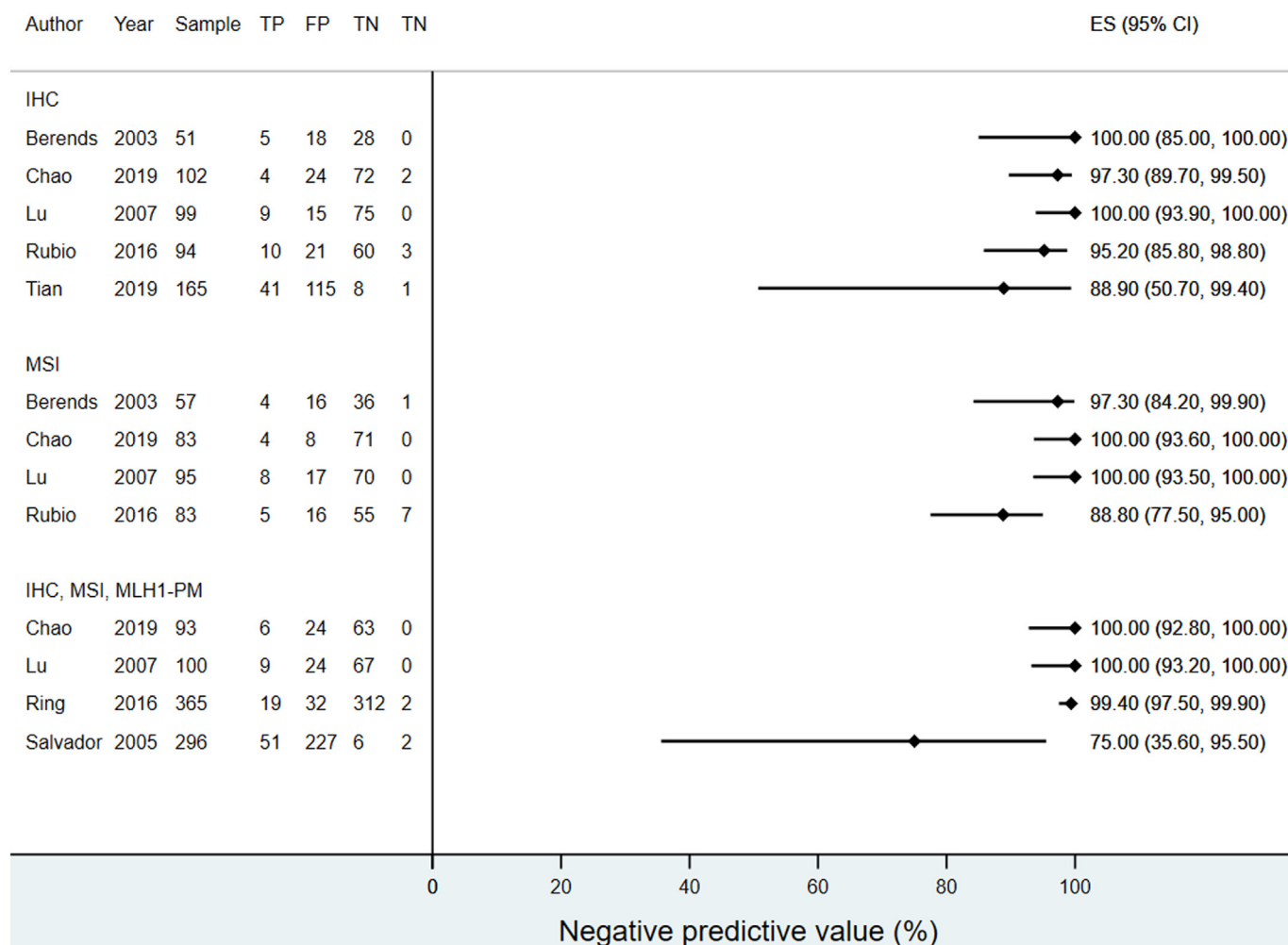


Fig. 4 (continued).

each protein separately [27]. In the study that combined results across proteins, there were 16 true positives, and 115 false positives from 500 women tested. PPV was 12.2% (95% CI 7.4–19.4%). In the study that reported results by genes, PPV was lowest for MLH1 (75.0%) and highest for PMS2 (81.8%) [27]. Test failures were reported in one study [16], in which 13 (2.6%) tumours failed initial sample testing. Repeat testing was successful for all 13 samples.

### 3.7.2. Immunohistochemistry with MLH1 promoter methylation testing

One study provided test accuracy data for immunohistochemistry with MLH1 promoter methylation testing [25]. There were 2 true positives, and 8 false positives from 23 women tested. PPV was 20% (95% CI 3.5–55.8%). No test failures were reported.

### 3.7.3. Microsatellite instability-based testing alone

Four studies provided test accuracy data for MSI [16,25,27,28]. There were 33 true positives and 197 false positives from 1072 women tested. PPV estimates ranged from 5.9–75%. Test failures were reported in one study [16], in which 8 (1.6%) tumours failed initial sample testing. Repeat testing was successful for all 8 samples.

### 3.7.4. Immunohistochemistry and microsatellite instability-based testing, with MLH1 promoter methylation testing

Three studies provided test accuracy data for IHC, MSI and MLH1 promoter methylation testing [24–26]. There were 28 true positive

and 46 false positives from 1138 women tested. PPV estimates ranged from 20 to 43.1%. Test failures were reported for 0–0.3% of tumours for IHC (3/1179 tumours), none for MSI testing, and 0–3.7% (40/1180 tumours) for MLH1 promoter hypermethylation testing.

## 4. Discussion

In this review, we evaluated the test accuracy of IHC and MSI (with or without MLH1 promoter methylation testing) for Lynch syndrome in women with endometrial cancer. In seven studies participants received both the index test and the reference standard (complete test accuracy studies), and in six studies only women who were positive on the index tests were followed up with the reference standard (partial test accuracy studies). In the complete test accuracy studies, sensitivity ranged from 60.7–100% for IHC, 41.7–100% for MSI, and 90.5–100% for IHC, MSI, and MLH1 promoter methylation testing combined. Specificity ranged from 6.5–83.3% for IHC, 69.2–89.9% for MSI, and 2.6–90.7% for IHC, MSI, and MLH1 promoter methylation testing combined. Two studies reported very low specificity: 2.6% [22], and 6.5% [21]. In each of these studies, the majority of participants had loss of MMR expression, due to highly selective inclusion criteria. Removing these studies gave specificity ranging from 60.9–83.3% for IHC, 69.2–89.9% for MSI, and 72.4–90.7% for IHC, MSI, and MLH1 promoter methylation testing combined.

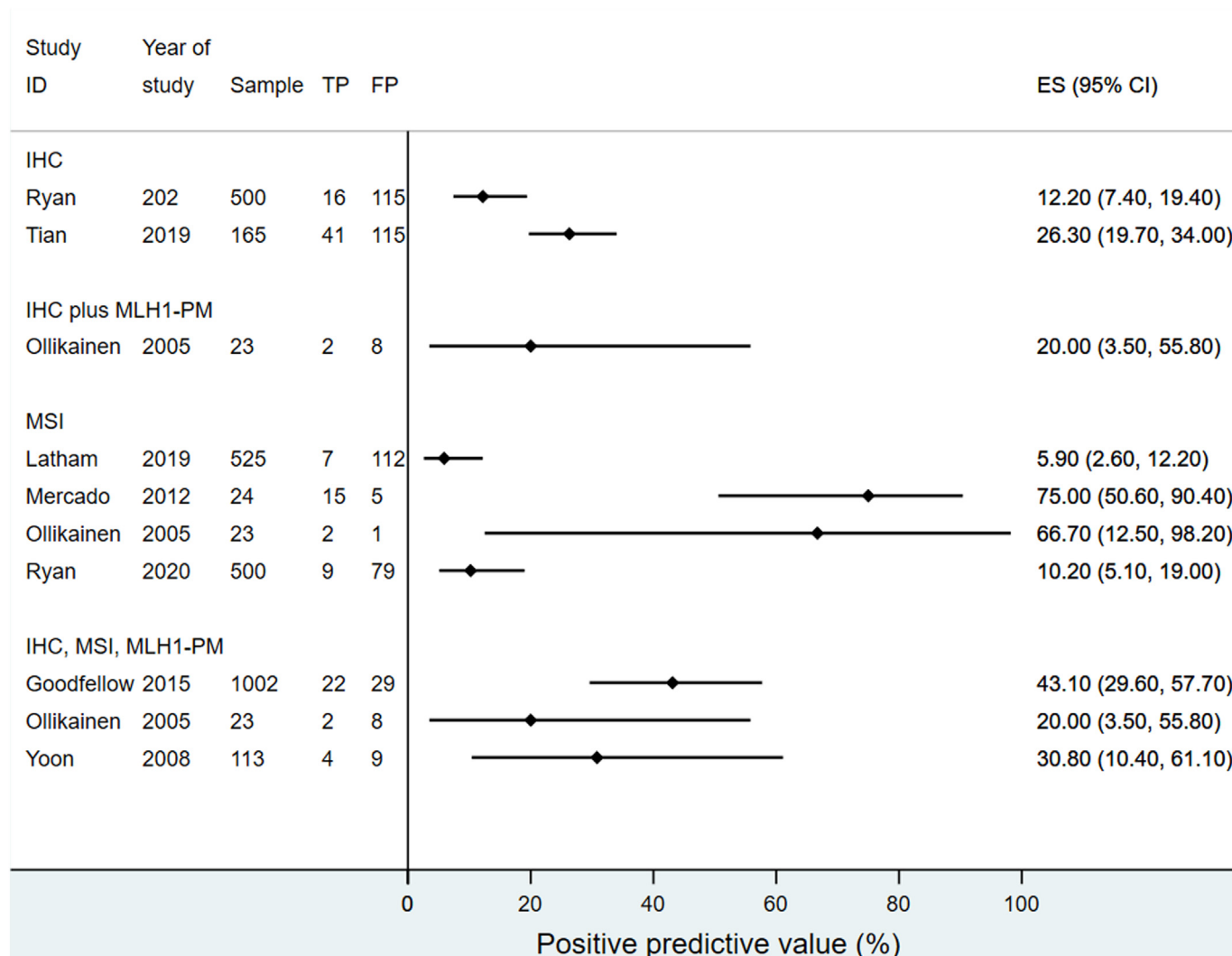


Fig. 5. Forest plots of partial test accuracy studies.

In the partial test accuracy studies, PPV estimates was 12.2 for IHC, and ranged from 5.9–75% for MSI, 20% for IHC with MLH1 promoter methylation testing, and 20–43.1% for studies that combined all three tests. There is a range of possible explanations for the differences in test accuracy estimates, including whether the studies included selected- (e.g. limiting participants by age, or prior/current cancer diagnosis) or unselected samples of participants, and variation in testing strategies (e.g. variable index test thresholds, the number of proteins assessed for IHC, the panel of markers used for MSI, and the reference standards used). We found no evidence of a difference in test accuracy between IHC and MSI in the four studies that provided head-to-head data. The number of tumours assessed by IHC and MSI in each studies were not identical. Typically, there were more IHC tests conducted than MSI tests. Therefore, the results of the comparisons are subject to bias. Test failures were extremely low for all three tests. This is possibly due to participants with insufficient tumour tissue available for testing being excluded from the studies.

Caution is warranted in drawing strong conclusions on test accuracy as the number of studies providing data on each test was small, as were the overall sample sizes and number of cases of Lynch syndrome. Further, the majority of studies were at high risk of bias in at least one domain, and there was considerable variability in study designs. Due to these and other issues (e.g. a lack of information on ideal marker panels for MSI-based strategies, and the impact of adding MLH1 promoter hypermethylation testing), it is currently unclear which method should be selected for clinical practice.

In our review we separated papers into ‘complete’ test accuracy studies (in which the reference standard was offered irrespective of the result of the index test) and ‘partial’ test accuracy studies (in which the reference standard was only offered to women who were index test positive. The rationale being that although in some studies index test negative cases are occasionally treated as true negatives to allow sensitivity and specificity to be calculated, this can lead to overestimation of sensitivity. Our study supports this. Using the seven complete test accuracy studies in our review, an assumption that the index test results were correct would have led to overestimation of sensitivity in 2/4 studies (50%) that assessed IHC alone (overestimating by 2.4–33.3%), 3/5 studies (60%) that assessed MSI alone (overestimating by 20–58.5%), and 2/4 studies (50%) that assessed IHC, MSI, and MLH1 promoter methylation testing combined (overestimating by 3.8–9.5%).

Two previous systematic reviews have examined the test accuracy of MSI [29], and IHC and MSI (with/without MLH1 promoter methylation testing) to detect Lynch syndrome, albeit in people who have colorectal cancer [11]. The authors reported sensitivity of 73.3–100% and specificity of 12.5–100% for IHC [11], and sensitivity of 66.7–100% and specificity of 61.1–92.5% [11,29]. These two reviews suggest that IHC and MSI test have similar levels of test accuracy to detect Lynch syndrome in people with colorectal and endometrial cancers. The low specificity reported in two studies suggests refinement of the testing strategies might be warranted before they are implemented in universal testing programmes, due to the cost implications and psychological impact of



testing people who are ultimately determined not to have Lynch syndrome.

The key strength of this review is that we followed the gold standard methodology for conducting systematic reviews (which included independent assessment at every stage) to identify evidence on test accuracy. Our review has a number of limitations. First, we excluded six studies where we could not establish which reference standard was used. In each case, we contact the paper's corresponding author. None of the authors who replied was able to confirm how the Lynch syndrome diagnosis was made, as study samples had been sent to commercial laboratories for assessment (in some cases more than one laboratory). Follow up with the relevant laboratories did not lead to confirmation of the tests used. Therefore, we cannot be certain these studies did not meet our inclusion criteria. Second, in our protocol we specified that Lynch syndrome had to be diagnosed by genetic confirmation of variants in the four MMR genes that are most commonly associated with Lynch syndrome (MLH1, MSH2, MSH6, and PMS2) using diagnostic tests that we described in the best practice guidelines of the Association for Clinical Genomic Sciences [30]. There is some evidence that another gene (EPCAM) is responsible for 1–3% of cases of Lynch syndrome [31]. Studies that were only concerned with the EPCAM gene, or that used diagnostic test not specified in the Association for Clinical Genomic Sciences guidelines would not have been captured in our review. Third, we excluded grey literature and studies not written in English. These studies might have provided additional information that was relevant to our review. Fourth, we excluded studies where fewer than 95% of participants who were eligible for the reference standard received it. In complete test accuracy studies, this referred to 95% of participants who had received the index test, and in the partial test accuracy studies, this referred to 95% of participants who were positive on the index test. We employed this approach to minimise biases in the test accuracy estimates, as systematic reasons why participants declined/were not offered the reference standard (e.g. people who are at higher/lower risk of a disease are systematically not being followed up) can affect test accuracy estimates. However, the decision to set the threshold at 95% was pragmatic.

## 5. Conclusions

Our review suggests that the sensitivity of the index tests were generally high. Most studies reported much lower levels of specificity, which might be problematic for universal screening programmes in terms of the costs and psychological implications of conducting genetic testing on women who do not have Lynch syndrome. We found no evidence that test accuracy differs between IHC and MSI based strategies. The evidence base is currently small and at high risk of bias. Additional studies that follow up participants who are index test negative with reference standards would provide greater clarity on the accuracy of IHC and MSI to detect Lynch syndrome in women who have endometrial cancer.

## Definition of test accuracy terms

False negative	The index test is negative but the person does have the disease
False positive	The index test is positive but the person does not have the disease
Negative predictive value	The probability that a person with a negative index test result truly does not have the disease
Positive predictive value	The probability that a person with a positive index test result truly does have the disease
Sensitivity	The ability of an index test to correctly identify people with the disease
Specificity	The ability of an index test to correctly identify people without the disease
True negative	The index test is negative and the person does not have the disease
True positive	The index test is positive and the person does have the disease

## Declarations

Ethical approval and consent to participate not applicable.

## Consent for publication

Not applicable.

## Availability of data and materials

The datasets supporting the conclusions of this article are included within the article and its additional files.

## Funding

This report was commissioned by the National Institute for Health and Care Excellence Diagnostic Assessment Programme as project number NIHR 129546. Sian Taylor-Phillips is funded by a NIHR Career Development fellowship (reference CDF-2016-09-018). The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

## Author contributions

Chris Stinton: Conceptualization; Methodology; Formal analysis; Investigation; Data Curation; Roles/Writing - original draft; Visualization. Hannah Fraser: Methodology; Investigation; Data Curation; Roles/Writing - original draft. Lena Al-Khudairy: Methodology; Investigation; Data Curation; Writing - review & editing. Rachel Court: Methodology; Investigation; Writing - review & editing. Mary Jordan: Methodology; Investigation; Writing - review & editing. Dimitris Grammatopoulos: Methodology; Conceptualization; Writing - review & editing. Sian Taylor-Phillips: Methodology; Conceptualization; Writing - review & editing; Supervision; Funding acquisition.

## Ethics statement

This article does not contain any studies with human or animal subjects performed by any of the authors.

## Declaration of Competing Interest

The authors declare that they have no competing interests.

## Acknowledgements

The authors would like to thank Wendy Knerr, Alexander Tsertsvadze, and Karoline Munro for assistance with data extraction. This project was funded via the National Institute for Health Research (grant number 14/25/05). Sian Taylor-Phillips is funded by an NIHR Career Development fellowship.

## Appendix A. Supplementary Data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ygyno.2020.10.003>.

## References

- [1] M.A. Roett, Genital cancers in women: uterine Cancer, *FP Essent.* 438 (2015) 11–17.
- [2] Uterine Cancer Statistics, <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/uterine-cancer#heading=Three> 2016.
- [3] Cancer Survival in England: National Estimates for Patients Followed up to 2017, <https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/conditionsanddiseases/bulletins/cancersurvivalinengland/nationalestimatesforpatientsfollowedupto2017> 2017.

- [4] O. Raglan, I. Kalliala, G. Markozannes, S. Cividini, M.J. Gunter, J. Nautiyal, H. Gabra, E. Parakevaidis, P. Martin-Hirsch, K.K. Tsilidis, et al., Risk factors for endometrial cancer: an umbrella review of the literature, *Int. J. Cancer* 145 (7) (2019) 1719–1730.
- [5] A.T. Ali, Reproductive factors and the risk of endometrial cancer, *Int. J. Gynecol. Cancer* 24 (3) (2014) 384–393.
- [6] B. Trabert, N. Wentzensen, H.P. Yang, M.E. Sherman, A.R. Hollenbeck, Y. Park, L.A. Brinton, Is estrogen plus progestin menopausal hormone therapy safe with respect to endometrial cancer risk? *Int. J. Cancer* 132 (2) (2013) 417–426.
- [7] P.M. Boland, M.B. Yurgelun, C.R. Boland, Recent progress in lynch syndrome and other familial colorectal cancer syndromes, *CA Cancer J. Clin.* 68 (3) (2018) 217–231.
- [8] P. Møller, T.T. Seppälä, I. Bernstein, E. Holinski-Feder, P. Sala, D. Gareth Evans, A. Lindblom, F. Macrae, I. Blanco, R.H. Sijmons, et al., Cancer risk and survival in path\_MMR carriers by gene and gender up to 75 years of age: a report from the prospective lynch syndrome database, *Gut* 67 (7) (2018) 1306–1316.
- [9] J.M. Lancaster, C.B. Powell, L.M. Chen, D.L. Richardson, Society of gynecologic oncology statement on risk assessment for inherited gynecologic cancer predispositions, *Gynecol. Oncol.* 136 (1) (2015) 3–7.
- [10] E.J. Crosbie, N.A.J. Ryan, M.J. Arends, T. Bosse, J. Burn, J.M. Cornes, R. Crawford, D. Eccles, I.M. Frayling, S. Ghaem-Maghami, et al., The manchester international consensus group recommendations for the management of gynecological cancers in lynch syndrome, *Genet. Med.* 21 (10) (2019) 2390–2400.
- [11] T. Snowsill, N. Huxley, M. Hoyle, T. Jones-Hughes, H. Coelho, C. Cooper, I. Frayling, C. Hyde, A systematic review and economic evaluation of diagnostic strategies for lynch syndrome, *Health Technol. Assess.* 18 (58) (2014) 1–406.
- [12] P.F. Whiting, A.W. Rutjes, M.E. Westwood, S. Mallett, J.J. Deeks, J.B. Reitsma, M.M. Leeflang, J.A. Sterne, P.M. Bossuyt, QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies, *Ann. Intern. Med.* 155 (8) (2011) 529–536.
- [13] P.F. Whiting, A.W. Rutjes, M.E. Westwood, S. Mallett, A systematic review classifies sources of bias and variation in diagnostic test accuracy studies, *J. Clin. Epidemiol.* 66 (10) (2013) 1093–1104.
- [14] E.B. Wilson, Probable inference, the law of succession, and statistical inference, *J. Am. Stat. Assoc.* 22 (158) (1927) 209–212.
- [15] Stata: Stata Statistical Software: Release 16., StataCorp LLC., College Station, TX, 2019.
- [16] N. Ryan, R. McMahon, S. Tobi, T.M. Snowsill, S. Esquibel, A.J. Wallace, S. Bunstone, N. Bowers, I.E. Mosneag, S. Kitson, et al., The proportion of endometrial tumours associated with lynch syndrome (PETALS): a prospective cross-sectional study, *PLOS Med.* 17 (9) (2020) e1003263.
- [17] M.J. Berends, Y. Wu, R.H. Sijmons, T. van der Sluis, W.B. Ek, M.J. Ligtenberg, N.J. Arts, K.A. ten Hoor, J.H. Kleibeuker, E.G. de Vries, et al., Toward new strategies to select young endometrial cancer patients for mismatch repair gene mutation analysis, *J. Clin. Oncol.* 21 (23) (2003) 4364–4370.
- [18] X. Chao, L. Li, M. Wu, S. Ma, X. Tan, S. Zhong, Y. Bi, J. Lang, Comparison of screening strategies for lynch syndrome in patients with newly diagnosed endometrial cancer: a prospective cohort study in China, *Cancer Commun. (Lond.)* 39 (1) (2019) 42.
- [19] K.L. Ring, A.S. Bruegl, B.A. Allen, E.P. Elkin, N. Singh, A.R. Hartman, M.S. Daniels, R.R. Broaddus, Germline multi-gene hereditary cancer panel testing in an unselected endometrial cancer cohort, *Mod. Pathol.* 29 (11) (2016) 1381–1389.
- [20] I. Rubio, E. Ibanez-Feijoo, L. Andres, E. Aguirre, J. Balmana, P. Blay, G. Lloret, S. Gonzalez-Santiago, H. Maortua, M.I. Tejada, et al., Analysis of lynch syndrome mismatch repair genes in women with endometrial cancer, *Oncology* 91 (3) (2016) 171–176.
- [21] W. Tian, R. Bi, Y. Ren, H. He, S. Shi, B. Shan, W. Yang, Q. Wang, H. Wang, Screening for hereditary cancers in patients with endometrial cancer reveals a high frequency of germline mutations in cancer predisposition genes, *Int. J. Cancer* 145 (5) (2019) 1290–1298.
- [22] M.U. Salvador, M.R.F. Truelson, C. Mason, B. Souders, H. LaDuca, B. Dougall, M.H. Black, K. Fulk, J. Profato, S. Gutierrez, et al., Comprehensive paired tumor/germline testing for lynch syndrome: bringing resolution to the diagnostic process, *J. Clin. Oncol.* 37 (8) (2019) 647–657.
- [23] K.H. Lu, J.O. Schorge, K.J. Rodabaugh, M.S. Daniels, C.C. Sun, P.T. Soliman, K.G. White, R. Luthra, D.M. Gershenson, R.R. Broaddus, Prospective determination of prevalence of lynch syndrome in young women with endometrial cancer, *J. Clin. Oncol.* 25 (33) (2007) 5158–5164.
- [24] S.N. Yoon, J.L. Ku, Y.K. Shin, K.H. Kim, J.S. Choi, E.J. Jang, H.C. Park, D.W. Kim, M.A. Kim, W.H. Kim, et al., Hereditary nonpolyposis colorectal cancer in endometrial cancer patients, *Int. J. Cancer* 122 (5) (2008) 1077–1081.
- [25] M. Ollikainen, W.M. Abdel-Rahman, A.L. Moiso, A. Lindroos, R. Kariola, I. Jarvela, M. Poyhonen, R. Butzow, P. Peltomaki, Molecular analysis of familial endometrial carcinoma: a manifestation of hereditary nonpolyposis colorectal cancer or a separate syndrome? *J. Clin. Oncol.* 23 (21) (2005) 4609–4616.
- [26] P.J. Goodfellow, C.C. Billingsley, H.A. Lankes, S. Ali, D.E. Cohn, R.J. Broaddus, N. Ramirez, C.C. Pritchard, H. Hampel, A.S. Chassen, et al., Combined microsatellite instability, MLH1 methylation analysis, and immunohistochemistry for lynch syndrome screening in endometrial cancers from GOG210: an NRG oncology and gynecologic oncology group study, *J. Clin. Oncol.* 33 (36) (2015) 4301–4308.
- [27] R.C. Mercado, H. Hampel, F. Kastanos, E. Steyerberg, J. Balmana, E. Stoffel, D.E. Cohn, F.J. Backes, J.L. Hopper, M.A. Jenkins, et al., Performance of PREMM(1,2,6), MMRpredict, and MMRpro in detecting lynch syndrome among endometrial cancer cases, *Genet. Med.* 14 (7) (2012) 670–680.
- [28] A. Latham, P. Srinivasan, Y. Kemel, J. Shia, C. Bandlamudi, D. Mandelker, S. Middha, J. Hechtman, A. Zehir, M. Dubard-Gault, et al., Microsatellite instability is associated with the presence of lynch syndrome pan-cancer, *J. Clin. Oncol.* 37 (4) (2019) 286–295.
- [29] H. Coelho, T. Jones-Hughes, T. Snowsill, S. Briscoe, N. Huxley, I.M. Frayling, C. Hyde, A systematic review of test accuracy studies evaluating molecular micro-satellite instability testing for the detection of individuals with lynch syndrome, *BMC Cancer* 17 (1) (2017) 836.
- [30] I. Frayling, I. Berry, A. Wallace, S. Payne, G. Norbury, Association for Clinical Genomic Science Best Practice Guidelines for Genetic Testing and Diagnosis of Lynch Syndrome, Association for Clinical Genomic Science, London, 2016.
- [31] K. Tuttlewska, J. Lubinski, G. Kurzawski, Germline deletions in the EPCAM gene as a cause of lynch syndrome - literature review, *Hered. Cancer Clin. Pract.* 11 (1) (2013) 9.