A systematic review and economic evaluation of diagnostic strategies for Lynch syndrome

Tristan Snowsill,[1]* Nicola Huxley,[1] Martin Hoyle,[1]
Tracey Jones-Hughes,[1] Helen Coelho,[1] Chris Cooper,[1]
Ian Frayling[2] and Chris Hyde[1]

[1]Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter, UK
[2]Institute of Medical Genetics, Cardiff University, Cardiff, UK

*Corresponding author

Declared competing interests of authors: none

Scientific summary

Scientific summary

Background

- Lynch syndrome (LS), previously known as hereditary non-polyposis colorectal cancer (HNPCC), is an inherited autosomal dominant disorder characterised by an increased risk of colorectal cancer (CRC) and cancers of the endometrium, ovary, stomach, small intestine, hepatobiliary tract, urinary tract, brain and skin, among others. The lifetime cancer risk is highest for CRC (range 22–82% by age 70 years).
- LS accounts for 0.3–2.4% of CRC, with a general population prevalence of 1 : 3100. It is caused by mutations in deoxyribonucleic acid (DNA) mismatch repair (MMR) genes, specifically MutL homologue 1 (*MLH1*), MutS homologues 2 (*MSH2*) and 6 (*MSH6*), and postmeiotic segregation increased 2 (*PMS2*). Loss of MMR proficiency in a cell leads to an inability to repair DNA mismatches and the proliferation of genetic mutations. These mutations are more likely in repetitive DNA sequences known as microsatellites, a phenomenon known as microsatellite instability (MSI).
- Identification of family members carrying a MMR gene defect is desirable, in order to offer colonoscopic surveillance and prophylactic surgery as appropriate.
- If LS is identified, biennial colonic surveillance commencing at 25 years is recommended. Surveillance should cease for individuals testing negative for a characterised pathogenic germline mutation present in family members.
- Currently, clinical criteria [Amsterdam criteria (AC) II or Revised Bethesda criteria] are used to assist with the diagnosis of LS. Laboratory techniques are also available, including testing tumour tissue using immunohistochemistry (IHC), MSI testing (now included in the Revised Bethesda criteria) and genetic testing for MMR mutations. Supplementary tests include *BRAF* V600E and methylation of *MLH1*.

  - MSI testing involves identifying reference markers. Tumours with no instability in any of the markers are considered microsatellite stable. Those with one, or more than one, mutated reference marker are considered to have low MSI or high MSI respectively (in the case of a five marker panel).
  - IHC is performed on MLH1, MSH2, MSH6 and PMS2 proteins. Negative staining indicates a mutation in the corresponding MMR gene, thus identifying the gene(s) most likely to harbour a mutation.
  - A limitation of IHC and MSI testing is the existence of *MLH1* silencing in approximately 15% of sporadic CRC cases, leading to a false-positive LS result.
  - Multiple methods have been used for constitutional genetic testing in LS. Multiplex ligation-dependent probe amplification is the preferred technique in the UK.

Objective

i. To determine the accuracy of tests for LS in all newly diagnosed persons with CRC < 50 years of age, and those considered according to clinical criteria to be at high risk.
ii. To determine the diagnostic utility and cost-effectiveness of genetic testing for LS in all newly diagnosed persons with CRC < 50 years of age, and those of strategies to test their close relatives.

Methods

Test accuracy systematic review

- The assessment comprises a systematic review of evidence on the accuracy of LS laboratory tests. A literature search was conducted on 30 April 2012 in a range of electronic databases [including MEDLINE (1946 to April week 3, 2012), EMBASE (1980 to week 17, 2012) and The Cochrane Library (inception to 30 April 2012)] and in trial registries. The European Medicines Agency website and Google were also searched.
- Studies were included if:

  - the persons presenting with CRC were < 50 years of age, considered at risk of LS or close relatives of individuals with proven LS
  - they compared tumour-based tests against constitutional genetic testing
  - the outcome related to diagnostic accuracy, for example sensitivity and specificity.

- No study design was excluded unless evidence on the test was already available from higher-level study designs.
- Data extraction and critical appraisal [using Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2)] was performed by two reviewers. Individual results were summarised in tables and text. Data pooling was not possible due to study heterogeneity.

Cost-effectiveness systematic review

- This aimed to review cost-effectiveness studies related to the identification and management of individuals with LS. A literature search was conducted on 29 February 2012 (updated 5 February 2013) in a range of electronic databases including MEDLINE (1946 to February week 3, 2012; updated search 1946 to January week 4, 2013), EMBASE (1980 to week 8, 2012; updated search 1980 to week 5, 2013) and The Cochrane Library (inception to 29 February 2012; updated search, inception to Issue 1 of 12, January 2013). Studies were included where:

  - the population was persons who have or may have LS
  - the intervention was a strategy or strategies that identify and/or manage LS in a given population
  - the comparator was current clinical practice
  - outcomes included costs or clinically relevant outcomes [e.g. life-years or quality-adjusted life-years (QALYs) gained, CRCs prevented, mutations detected]
  - the study design was a decision-analytic model, evaluation of cost-effectiveness within trials, cost or resource use study, or guideline from a national institution or a professional or international body.

- Data extraction was carried out by two reviewers. Included studies were assessed for quality using the Drummond checklist. Data were synthesised using tables and text.

Peninsula Technology Assessment Group cost-effectiveness analysis
Our model of the cost-effectiveness of systematic screening for LS comprises a diagnostic and a survival component.

Diagnostic model
This is a decision tree model of short-term outcomes of diagnosis in probands and relatives.

We considered the following strategies to identify LS in probands:

1. No genetic testing, subdivided:

   1(1) no testing
   1(2) AC II.

2. IHC four-panel test, followed by genetic testing.
3. IHC, followed by *BRAF* testing then genetic testing.
4. MSI testing followed by genetic testing.
5. MSI testing, followed by *BRAF* testing, followed by genetic testing.
6. MSI testing, followed by *BRAF* testing, followed by IHC testing, followed by genetic testing.
7. IHC testing, followed by genetic testing if result abnormal. For normal IHC results: MSI testing, followed by *BRAF* testing for MSI result, followed by genetic testing for negative *BRAF* test.
8. Universal genetic testing.

The diagnosis of LS in relatives of a newly diagnosed CRC proband directly depended on the diagnosis of the proband, and predictive genetic testing was used when applicable.

A proportion of probands and relatives diagnosed with LS were assumed to undertake biennial surveillance colonoscopy and prophylactic total abdominal hysterectomy with bilateral salpingo-oophorectomy (TAHBSO).

The prevalence of LS and the sensitivities and specificities of individual tests were taken from published literature. Acceptance of tests was primarily based on expert opinion. Numbers of probands and relatives were taken from UK sources (Office for National Statistics data, published studies and unpublished data). The costs of the preliminary tumour tests and genetic tests were obtained directly from laboratories in the UK or from experts. The costs of genetic counselling and family history assessment were estimated using the Personal Social Services Research Unit and expert advice.

The psychological impact of testing for LS and prophylactic TAHBSO were incorporated into overall health-related quality of life (HRQoL), using data from the literature.

Survival model

This uses an individual patient simulation of thousands of hypothetical patients from time of LS diagnosis to death (or age 100 years). For each person, total costs and QALYs were calculated using methodology recommended by the National Institute for Health and Care Excellence, with costs and benefits discounted at 3.5% per annum. The model only considers the risks of CRC and endometrial cancer (EC).

Patient state at any time is defined by the following characteristics: age, sex, EC/CRC status, previous surgery (bowel or TAHBSO), LS status and diagnosis, acceptance of LS surveillance and whether or not the patient is alive.

Age at entry is a function of sex, true LS status and whether proband or relative. In the base-case analysis, the maximum age of probands is 50 years.

Simulated clinical events included incidence of CRC and EC; surgery for CRC and EC; colonoscopies (including bleeding and perforation); and mortality from CRC, EC, colonoscopy and background causes. The events determine costs incurred and HRQoL for each simulated patient. These are used to estimate the total discounted costs and QALYs for each testing strategy.

Parameters of the natural histories of diseases, the effectiveness of interventions and the impact on quality of life of diseases and interventions were sourced, where possible, from national statistics and published literature.

Costs of interventions were estimated from Department of Health reference costs 2011–12 with inflation to 2013–14 prices, or from published literature with appropriate conversion. The cost of a colonoscopy was adjusted to allow for the fact that the effectiveness of colonoscopy was taken from a regime of 3-yearly colonoscopy.

Uncertainty

We investigated uncertainty using scenario analyses and univariate sensitivity analyses upon the majority of parameters.

Results

Test accuracy systematic review

- Ten published papers were included (nine test accuracy studies and one technology assessment (TA) commissioned by the US Department of Health and Human Services).
- The TA found minimal published information on the analytical validity of laboratory testing for LS. Results ranged from 18% to 100% for sensitivity and 25% to 100% for specificity, with wide confidence intervals.

  ○ Many primary studies recruited preselected patients (e.g. from registries or pre-tests). However, those studies recruiting from a population that had no prior testing may include an increased number of false positives (FPs) due to *MLH1* methylation found in sporadic CRC. Other issues include: the reference standard was often not performed on all patients; sample sizes were generally small; and details on patient characteristics and robustness of testing were often lacking.

- Owing to the range of study designs, pooling of data was not possible.
- IHC sensitivity ranged from 73.3% to 100.0% and specificity from 12.5% to 100.0%. Specificity is the greatest concern; a high number of FPs means that individuals may be told they have LS when they do not.
- MSI sensitivity ranged from 88% to 100% and specificity from 68% to 84%. However, no two included studies used the same panel of markers.

Cost-effectiveness systematic review

- Thirty-two separate studies were identified, which examined strategies only identifying LS (15 studies); strategies only managing patients with LS (four studies); and strategies to both identify and manage LS (13 studies).
- The studies that included diagnosis and management were most relevant to our assessment. None of these were UK studies. Populations, settings and diagnostic strategies varied across the studies, and most only considered CRC in the long term. Quality assessment found that one consistent problem was the reporting of study viewpoint. Depth of detail related to modelling was mixed and, in particular, the justification for ranges of values in the sensitivity analyses was poorly reported. Study design was predominantly decision modelling. Most studies reported life-years and costs as their main outcomes, with two explicitly modelling QALYs.
- Generally, strategies that identified LS were found to be cost-effective compared with no LS screening. There was little consistency in terms of which strategies were the most cost-effective.

Peninsula Technology Assessment Group cost-effectiveness analysis

Base-case results

- Life expectancy of probands and relatives with LS improves by up to 1.6 years with testing.
- The expected total number of colonoscopies performed for probands aged < 50 years and their relatives in England, per year, increased from approximately 4200 in those given no testing to 8600 in strategy 8.
- The expected number of new CRC cases for the entire cohort in England, per year, reduces by up to 32 with testing.
- The expected annual number of ECs in England is reduced by up to nine with testing.
- Incremental cost-effectiveness ratios (ICERs) (vs. no testing) varied from £5491 per QALY for strategy 5 to £9571 per QALY for strategy 8.
- The testing strategies on the efficiency frontier were strategies 1(1), 5, 7 and 8. The remaining strategies were either dominated (less effective and more expensive than at least one other strategy) or extended dominated (less effective and more expensive than some combination of two other strategies). On the efficiency frontier, the ICER of strategy 5 versus no testing was £5491 per QALY. The ICER of strategy 7 versus strategy 5 was £25,106 and the ICER of strategy 8 versus strategy 7 was £82,962 per QALY.
- At a willingness-to-pay threshold of £20,000 per QALY, strategies 4, 5, 6 and 7 offered the best value for money, with similar cost-effectiveness. These strategies are predicted to result in an additional 130 discounted QALYs per year (or the total discounted QALYs accrued over the lives of approximately five people) in England compared with no testing.

Increasing the maximum age of probands

- When the age limit for proband testing was raised to 60 or 70 years, strategies became worse value for money versus no testing compared with the base case. At the age limit of 60 years, all ICERs compared with no testing remained below the £20,000-per-QALY threshold, but at age 70 years the ICER for strategy 8 was above the £20,000-per-QALY threshold.
- The incremental net health benefit (INHB) at the population level compared with no testing increased in most strategies compared with the base case. Strategy 5 gave the greatest INHB at a willingness-to-pay of £20,000 per QALY in both cases: 193 discounted QALYs for the population of England per year when the age limit was 60 years, and 271 discounted QALYs when the age limit was 70 years.

Endometrial cancer excluded

- This scenario resulted in reduced costs and slight increase in life expectancy (therefore reduced ICERs), plus no disutility from EC, compared with base case. Thus, all strategies became more cost-effective compared with no testing. The ranking of cost-effectiveness among strategies remained the same.

BRAF replaced by methylation testing

- When *BRAF* testing was replaced by methylation testing in strategies 3, 5, 6 and 7, their cost-effectiveness changed marginally.
- The INHB of all four strategies decreased versus no testing at a threshold of £20,000 per QALY.

Univariate sensitivity analyses

- Several univariate sensitivity analyses were conducted to investigate the impact of various parameters on the cost-effectiveness results. Incidence of CRC for individuals with LS, mean number of relatives identified per proband, hazard ratio for colonoscopy in the prevention of metachronous CRC, cost of colonoscopy and length of time of psychological disutilities all had a substantial impact on cost-effectiveness, but the testing strategies all remained cost-effective at a threshold of £20,000 per QALY.
- When a disutility of 0.1 for prophylactic TAHBSO was assumed, all strategies resulted in greater costs and reduced QALYs compared with no testing.

Suggested research priorities

We recommend further research as follows:

- Model the cost-effectiveness of testing for LS in probands newly diagnosed with EC and, separately, probands presenting with ovarian cancer and perhaps rarer LS-associated cancers.
- Incorporate aspirin (CRC prevention) in the model.
- Investigate disutilities for patients with CRC and disutilities after TAHBSO, particularly because the cost-effectiveness of genetic testing is very sensitive to the latter.
- Research the psychological impact of genetic testing for LS on HRQoL. The current evidence is extremely weak.
- Investigate the accuracy of individual tests when they are performed in sequence after early tests, i.e. in enriched populations.
- The cost-effectiveness model could be adapted for use in other countries.

Study registration

This study is registered as PROSPERO CRD42012002436.

Funding

Health Technology Assessment

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: http://www.nets.nihr.ac.uk/programmes/hta

This report