Iridology: a systematic review

Ernst E

Authors' objectives
To systematically review all interpretable tests of the validity of iridology as a diagnostic tool.

Searching
MEDLINE, EMBASE and CISCOM were searched from inception to 1997. In addition, experts in complementary medicine were asked for further references, professional iridology societies were invited to contribute material, and the author's own files were searched. All bibliographies of identified articles were checked. No language restrictions were applied. Only peer-reviewed articles were included.

Study selection

Controlled blinded evaluations of diagnostic validity were eligible. Case-control studies were included.

Specific interventions included in the review
Studies evaluating iris diagnosis were eligible. Three studies tested the ability of human observers to discriminate between photographic images of the irises of: (a) patients with specified conditions; and (b) controls without the conditions. A fourth study tested photodensitometry of colour images in an automated process that did not involve human observers.

Reference standard test against which the new test was compared
No inclusion criteria relating to the reference standard were specified. The included studies examined a variety of conditions. The method used to confirm diagnosis was outlined for two of the four included studies: a creatinine level above 1.6 mg/dL for kidney disease; and subsequent post-operative confirmation for inflamed gallbladder. The methods used to confirm diagnoses of mitral stenosis, ulcerative colitis, asthma, coronary heart disease or psoriasis, were not stated.

Participants included in the review
The inclusion criteria were not defined in terms of the participants. Each study included patients with previously diagnosed conditions and controls without the condition.

Outcomes assessed in the review
No inclusion criteria relating to the outcomes were stated. The included studies were designed to test whether iridology could discriminate between patients with previously diagnosed conditions and controls without the condition. The percentage validity, sensitivity and specificity were given for one study only; the results of the remaining studies were given as significance levels.

How were decisions on the relevance of primary studies made?
The author did not state how the papers were selected for the review, or how many reviewers performed the selection.

Assessment of study quality
Controlled evaluator-blinded studies were accepted. Validity was not formally assessed, although aspects of validity were discussed.

Data extraction
Data were extracted on the study characteristics, patients, controls, technique, observers and results. The author stated that this was done in a predefined, standardised fashion, but omitted any details.
Methods of synthesis
How were the studies combined?
The results of each study were reported separately and discussed in the text.

How were differences between studies investigated?
Differences between the studies were not discussed.

Results of the review
Four diagnostic studies were included. There was a discrepancy between the sample sizes reported in the data table (n=385) and the text (n=388); the data table also omitted the controls for one study.

In three studies iridology was stated to be no more discriminatory than chance. In one study of the diagnosis of mitral stenosis, photometric values of a defined area were found to be significantly different between the patients and controls (P<0.05).

Authors' conclusions
Iridology is not a useful diagnostic tool. It has the potential for doing damage in economic and health terms.

CRD commentary
The review asked a relevant question, although there were limited inclusion criteria and a very varied set of conditions. The literature search was conducted appropriately, but was not comprehensive, and the exclusion of non-peer-reviewed articles may have led to further publication bias. Details of the study selection, data extraction and checking processes were incomplete; since this was a single-author review, the possibility of reviewer error or bias cannot be ruled out. Predefined outcome criteria were also missing, and there was a particular lack of detail on the patients and the reference standards used for diagnosis.

There were noticeable errors in the reporting of basic sample data, e.g. the sample sizes included controls in three studies, but not in the mitral stenosis study. The reporting of the results was mostly incomplete: only one study reported the validity, sensitivity and specificity; a significant P-value was given for one study, and the other studies were merely stated to be not significant. There were also errors of interpretation: the single significant result (mitral stenosis) was dismissed for being based on a small sample (n=23), although the review omitted the controls from the sample size and a significant result is less likely from a small sample, not more likely. The same study was also dismissed for giving insufficient details of observer blinding, when it was the only study not to involve human observers.

The author's outright condemnation of iridology is questionable when the included studies did not contribute evidence either way, indicating that further research is needed. The review should not be regarded as reliable in view of problems relating to conduct, reporting accuracy and interpretation.

Implications of the review for practice and research
Practice: The author stated that patients and therapists should be discouraged from using this method.

Research: The author did not state any implications for further research.

Bibliographic details

PubMedID
10213874
Record Status
This is a critical abstract of a systematic review that meets the criteria for inclusion on DARE. Each critical abstract contains a brief summary of the review methods, results and conclusions followed by a detailed critical assessment on the reliability of the review and the conclusions drawn.