EBioMedicine 10 (2016) 164–173

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.ebiomedicine.com

Research Paper

Building the Evidence Base of Blood-Based Biomarkers for Early Detection of Cancer: A Rapid Systematic Mapping Review

Lesley Uttley [a,1], Becky L. Whiteman [b,c,1], Helen Buckley Woods [a], Susan Harnan [a], Sian Taylor Philips [c], Ian A. Cree [b,d,*], For the Early Cancer Detection Consortium

[a] The University of Sheffield, Regent Court, 30 Regent Street, Sheffield S1 4DA, UK
[b] Centre for Technology Enabled Health Research, Faculty of Health and Life Sciences, Coventry University, Coventry CV1 5FB, UK
[c] Warwick Medical School, University of Warwick, Coventry CV4 7AL, UK
[d] Department of Pathology, University Hospitals Coventry and Warwickshire, Coventry CV2 2DX, UK

ARTICLE INFO

Article history:
Received 13 January 2016
Received in revised form 24 June 2016
Accepted 5 July 2016
Available online 6 July 2016

Keywords:
Cancer
Early detection
Biomarker
Assay
Diagnosis
Blood
Systematic review

ABSTRACT

Background: The Early Cancer Detection Consortium is developing a blood-test to screen the general population for early identification of cancer, and has therefore conducted a systematic mapping review to identify blood-based biomarkers that could be used for early identification of cancer.

Methods: A mapping review with a systematic approach was performed to identify biomarkers and establish their state of development. Comprehensive searches of electronic databases Medline, Embase, CINAHL, the Cochrane library and Biosis were conducted in May 2014 to obtain relevant literature on blood-based biomarkers for cancer detection in humans. Screening of retrieved titles and abstracts was performed using an iterative sifting process known as "data mining". All blood based biomarkers, their relevant properties and characteristics, and their corresponding references were entered into an inclusive database for further scrutiny by the Consortium, and subsequent selection of biomarkers for rapid review. This systematic review is registered with PROSPERO (no. CRD42014010827).

Findings: The searches retrieved 19,724 records after duplicate removal. The data mining approach retrieved 3990 records (i.e. 20% of the original 19,724), which were considered for inclusion. A list of 814 potential blood-based biomarkers was generated from included studies. Clinical experts scrutinised the list to identify miss-classified and duplicate markers, also volunteering the names of biomarkers that may have been missed: no new markers were identified as a result. This resulted in a final list of 788 biomarkers.

Interpretation: This study is the first to systematically and comprehensively map blood biomarkers for early detection of cancer. Use of this rapid systematic mapping approach found a broad range of relevant biomarkers allowing an evidence-based approach to identification of promising biomarkers for development of a blood-based cancer screening test in the general population.

1. Introduction

Early detection of cancer results in improved survival (Etzioni et al., 2003; Wolf et al., 2010; McPhail et al., 2015). Cancers detected early require less extensive treatment and are less likely to have spread to other organs. Cancer diagnosis requires histological examination of tissue abnormalities detected by radiological, clinical or endoscopic examination of patients. Detection, as opposed to diagnosis, relies on screening a largely asymptomatic population to identify people who may be at higher risk of having cancer than others. Screening tests for cancer, or any other condition need to fulfil strict criteria to prevent the implementation of inappropriate screening, ensuring screening is cost effective and benefits patients. The criteria applied within the UK are listed at http://www.screening.nhs.uk/criteria, based on those developed by Wilson and Jungner (Cochrane & Holland, 1971; Wilson & Jungner, 1968). For early cancer detection, a blood-based screening test would have to be cost effective and demonstrate a meaningful clinical benefit which outweighs the harms associated with false positive, indeterminate results and overtreatment. This is clearly a major undertaking, and needs a multidisciplinary approach.

The Early Cancer Detection Consortium (ECDC) was established in 2012 in the United Kingdom and comprises 23 universities, their associated NHS hospitals, as well as other organisations and industry partners. The consortium was established to investigate whether a cost-effective screening test can be used in the general population to identify people

* Corresponding author at: Department of Pathology, University Hospitals Coventry and Warwickshire NHS Trust, Clifford Bridge Road, Coventry CV2 2DX, UK.
E-mail address: ian.cree@uhcw.nhs.uk (I.A. Cree).
[1] Joint first authorship.

http://dx.doi.org/10.1016/j.ebiom.2016.07.004

Table 1
Eligibility criteria for the systematic mapping review.

Inclusion criteria	Exclusion criteria
English language studies	Studies published in non-English language
Studies within the last five years (2010–2014)	Studies from 2009 or older
Controlled studies	No healthy control group
Validation studies	Derivative studies from included papers
Cancer detection/diagnosis	Prognosis or prediction (treatment response) associated markers
50 or more patients	Less than 50 patients
Biomarkers measured in blood	Tissue or other bodily fluid samples
	Abstracts of panels which do not state which biomarkers are studied
	Citation titles without abstracts

with early cancers. Given the extensive literature on blood biomarkers for cancer, it is logical to explore the development of such a test using existing biomarkers that have the best evidence-base for cancer detection. A sensitive blood test for multiple tumour types could enable people with biomarker levels which are outside the typical range to receive further investigation and lead to earlier diagnosis of cancer at an asymptomatic stage when curative treatment is feasible. The next stage of the programme will involve analytical and clinical validation of these biomarkers in a case control study, from which a detection algorithm will be produced and validated for possible use as a generic cancer screen. Finally, a randomised controlled trial will be required to determine the clinical and cost-effectiveness of the resulting screening strategy.

Previous reviews in this area have understandably been limited in scope, usually restricted to one biomarker or well-defined group of potential markers, due to the enormous number of publications in the field. The aim of this study was therefore to establish the full range of candidate blood-based biomarkers with potential for the early detection of cancer, and map key characteristics of the tests.

2. Methods

To identify all relevant biomarkers, comprehensive searches and innovative methods to perform the mapping review were employed to cope with the sizeable body of relevant literature to be assessed within a short time-frame. The mapping review comprised the following stages: comprehensive literature searches; data mining techniques for rapid screening of the search records and; development of a customizable database of evidence to optimise the output from the mapping review. It was not considered sufficient simply to list evidence by reference or to name the biomarker once in a spreadsheet and continue searching until another new biomarker was found. Instead it was more useful and time-efficient to maintain the corresponding citations for each biomarker and record the basic characteristics of the study at the time of screening. This enabled a basic informative profile to be built for each biomarker identified in the mapping review.

This systematic review is registered with PROSPERO (no. CRD42014010827) and the methods have been structured around the PRISMA checklist (http://www.prisma-statement.org/).

2.1. Eligibility Criteria

Eligible studies included all English language studies from the past five years that investigated blood based biomarkers in more than 50 patients, see Table 1.

2.2. Search Strategy

To identify a comprehensive body of literature from which a list of candidate biomarkers could be generated, a broad search using keywords and subject headings was undertaken. The terms reflected the concepts of 'diagnosis', 'markers', 'blood' and 'screening' (see supplementary material). The keywords and subject headings were developed using a variety of collaborative methods between Information Specialists and Systematic Reviewers at the University of Sheffield and researchers at the University of Warwick.

A scoping search was performed and assessed for appropriateness. Additionally, key journal articles and abstracts in Medline were retrieved and assessed to obtain relevant subject headings and keywords. Clinical input was sought from members of the ECDC to verify and validate the chosen keywords. For the full search, relevant free-text, keyword and thesaurus terms were combined using Boolean operators and translated into database specific syntax. Full searches were limited to English language, humans and publication dated from 2010 to May 2014. The databases searched were Medline and Medline in Process, Embase, CINAHL, Cochrane Library (including Cochrane Database of Systematic Reviews, DARE, CENTRAL, HTA, NHS EED), Science Citation Index Expanded, Conference Proceedings Citation Index - Science, Book Citation Index – Science, and Biosis Previews.

The initial search strategy was broad and inclusive. As a result, a large number of relevant records were obtained. Preliminary validation by consulting experts in the field indicated that the search was sensitive and no missing relevant literature was identified.

2.3. Sifting and Data Mining

The results of the initial searches were imported into a Reference Manager database. To identify an exhaustive list of biomarkers, retrieved records were searched iteratively within the Reference Manager database, using keywords to select potentially relevant titles. Titles and abstracts of this selection of citations were scrutinised for names and descriptions of biomarkers that met (or potentially met) the selection criteria (see Table 1). The citations were tagged to indicate that they had been viewed, to enable their exclusion from further searches. Relevant citations were exported to a Microsoft Access database which was customised to allow data extraction of relevant key information for each biomarker that was available from the corresponding study abstracts.

The data mining process within the main database included the following restrictions (see Box 1):

To ensure a comprehensive capture of all relevant biomarkers, a further validation stage was performed. Relevant reviews identified during the search were used to check for additional biomarkers not generated by the data mining process. ECDC members were invited to recommend papers that they believed to be relevant to the mapping review.

2.4. Data Collection

Each biomarker occupied a record with a unique identifier number in a customised Microsoft Access database which stored the number of associated papers, the abstract and reference details; associated synonyms and acronyms; types of cancers and study design; keywords used to retrieve the abstract during data mining; assays used to measure the biomarker, where reported; category to which the biomarker was assigned (e.g. auto-antibodies); and the sample types used, where reported (e.g. serum, plasma or whole blood).

2.5. Results

After duplicates were removed, 19,724 records were yielded from the comprehensive searches. Using data mining, 3990 titles and abstracts were retrieved from the 19,724 records for full scrutiny. Data mining is the process of pulling a subset of records from a large, unwieldy dataset. The subset of 3990 abstracts was reviewed in order to generate a list of biomarkers which are potentially relevant to early identification of cancer using blood. A full breakdown of the keywords used and the number of corresponding records retrieved can be seen in Fig. 1. During the validation process, three relevant reviews were