ELSEVIER

Journal of Clinical Epidemiology ■ (2016) ■

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Empirical evaluation demonstrated importance of validating biomarkers for early detection of cancer in screening settings to limit the number of false-positive findings

Hongda Chen[a], Phillip Knebel[b], Hermann Brenner[a,c,d,*]

[a]*Division of Clinical Epidemiology and Aging Research, German Cancer Research Center (DKFZ), Im Neuenheimer Feld 581, 69120 Heidelberg, Baden-Württemberg, Germany*

[b]*Department of General, Visceral and Transplantation Surgery, University of Heidelberg, Im Neuenheimer Feld 110, 69120 Heidelberg, Baden-Württemberg, Germany*

[c]*Division of Preventive Oncology, German Cancer Research Center (DKFZ) and National Center for Tumor Diseases (NCT), Im Neuenheimer Feld 460, 69120 Heidelberg, Baden-Württemberg, Germany*

[d]*German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, 69120 Heidelberg, Baden-Württemberg, Germany*

Accepted 25 January 2016; Published online xxxx

Abstract

Objectives: Search for biomarkers for early detection of cancer is a very active area of research, but most studies are done in clinical rather than screening settings. We aimed to empirically evaluate the role of study setting for early detection marker identification and validation.

Study Design and Setting: A panel of 92 candidate cancer protein markers was measured in 35 clinically identified colorectal cancer patients and 35 colorectal cancer patients identified at screening colonoscopy. For each case group, we selected 38 controls without colorectal neoplasms at screening colonoscopy. Single-, two- and three-marker combinations discriminating cases and controls were identified in each setting and subsequently validated in the alternative setting.

Results: In all scenarios, a higher number of predictive biomarkers were initially detected in the clinical setting, but a substantially lower proportion of identified biomarkers could subsequently be confirmed in the screening setting. Confirmation rates were 50.0%, 84.5%, and 74.2% for one-, two-, and three-marker algorithms identified in the screening setting and were 42.9%, 18.6%, and 25.7% for algorithms identified in the clinical setting.

Conclusion: Validation of early detection markers of cancer in a true screening setting is important to limit the number of false-positive findings.

Keywords: Research design; Biomarker identification; Early detection; Validity; Study setting; Methodology

1. Introduction

For most cancers, prognosis strongly varies by stage at diagnosis, and the prospect of cure is much higher when the cancer is detected at an early stage. Search for and validation of biomarkers for early detection of cancer is therefore a very active area of research [1−3]. Ideally, pertinent studies should be conducted in a true screening setting to provide reliable estimates of diagnostic performance in the target population for screening [4−6]. However, this is rarely done in practice for several reasons: First, the prevalence of preclinical cancer overall and of specific cancers in particular in the target population for cancer screening (which typically consists of essentially healthy older adults) is typically very low [7]. Therefore, very large study populations are required to ensure sufficient numbers of cases to estimate sensitivity and other indicators of diagnostic performance with adequate precision. Second, there is often no easy to perform and reliable gold standard

Funding: The BliTz study was partly funded by grants from the German Research Council (DFG, grant no. BR1704/16-1). The work of H.C. was partly supported by China Scholarship Council (CSC). The sponsor had no role in the study design, in the collection, analysis, and interpretation of data.

Conflict of interest: The authors disclose no potential conflicts of interest.

* Corresponding author. Tel./fax: +49-6221-42130.

E-mail address: h.brenner@dkfz-heidelberg.de (H. Brenner).

http://dx.doi.org/10.1016/j.jclinepi.2016.01.022

What is new?

Key findings

- Most studies searching for novel biomarkers for early detection of cancer are conducted in clinical settings, that is, using clinically detected cases. Biomarker levels among such cases may differ from biomarker levels among preclinical cases to be detected by screening for a variety of reasons. Biomarkers identified in such studies may therefore be of questionable use unless they are validated in a true screening setting.
- The authors provide a thorough quantitative illustration of the importance of validation of biomarkers for early detection of cancer in a true screening setting using the example of blood protein markers for early detection of colorectal cancer.

What is the implication and what should change now?

- The interpretation of case−control studies based on patients from clinical settings requires particular caution, due to the potential high proportion of false-positive findings.
- Validation of early detection biomarkers of cancer in a true screening setting is important to limit the number of false-positive findings.

examination (such as colonoscopy for colorectal cancer detection) to which measurements of biomarkers could be compared and that could be applied in such large screening populations.

In practice, it is commonly seen that studies evaluating biomarkers for early detection of cancer therefore recruit a sample of cancer cases in a clinical setting (e.g., newly diagnosed cancer patients admitted to a single or multiple clinics), along with a sample of controls without a known cancer diagnosis [8,9]. Moreover, controls often consist of convenience samples, such as patients from different clinic departments or healthy volunteers, and they often strongly differed from the case groups in many respects, including basic sociodemographic factors [10], such as presence of other diseases or age, which may have important implications for specificity. Under such circumstances, any differences in biomarker levels between cancer patients and controls need to be interpreted with caution because they could simply reflect such differences rather than cancer-related differences. In some studies, matching by key sociodemographic factors, such as sex and age, is used to reduce the risk of such bias, but such matching does not eliminate other sources of differences, such as preceding diagnostic measures that led to the diagnosis of cases or even early treatment. Furthermore, even seemingly perfect matching by factors such as sex and age may sometimes introduce or increase rather than eliminate bias because in a true screening setting, age and sex distribution of those with and without cancer is often not identical [11]. Finally, clinically manifest cases are by definition different from preclinical cancers searched for in a screening setting, and they may differ with respect to a number of factors that favor clinical diagnosis, such as cancer size or stage [10].

It is therefore not surprising that very promising results for the diagnostic performance of cancer early detection markers initially obtained in studies conducted in clinical settings could often not be confirmed in later validations in screening settings. On the other hand, it could be anticipated that good diagnostic performance in screening settings should more often go along with good diagnostic performance in clinical settings. This is because studies conducted in screening populations should primarily identify cancer-related differences (e.g., different expression patterns of tumor-associated biomarkers) between cases and controls which would also be expected to apply to clinical settings. However, evidence on differences in confirmation rates of early detection markers identified in clinical and screening settings from systematic comparative assessment is still sparse. In this study, we provide such an assessment, using the search for blood protein biomarkers for early detection of colorectal cancer as an example.

2. Methods

2.1. Study design

We compared the frequency of initial identification and subsequent validation of protein markers and protein marker combinations indicative of presence of colorectal cancer for the following two scenarios:

1) Use of clinically detected cases in the marker identification set and cases detected in a true screening setting in the validation set.
2) Use of cases detected in a true screening setting in the marker identification set and clinically detected cases in the validation set.

In both scenarios, two sets of participants confirmed to be free of colorectal neoplasms at screening colonoscopy were used as controls.

For this comparison, number and composition of study participants in the clinical setting and the screening setting were kept identical. To achieve statistically robust results, a large number of biomarkers and their combinations were evaluated: 92 single protein markers, $\binom{92}{2}=4,186$ two-marker combinations, and $\binom{92}{3}=125,580$ three-marker combinations.