Clinical effectiveness of cancer screening biomarker tests offered as self-pay health service: a systematic review

Agnes Luzak¹,² Petra Schnell-Inderst¹,², Stefanie Bühn¹,², Anja Mayer-Zitarosa¹,², Uwe Siebert¹,²,³,⁴

1 Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Hall i. T, Austria
2 Division of Health Technology Assessment and Bioinformatics, Oncoryol - Center for Personalized Cancer Medicine, Innsbruck, Austria
3 Center for Health Decision Science, Department of Health Policy and Management, Harvard T.H. Chan School of Public Health, Boston, MA, USA
4 Institute for Technology Assessment and Department of Radiology, Massachusetts General Hospital, Harvard Medical School, Boston, MA, USA

Correspondence: Uwe Siebert, Institute of Public Health, Medical Decision Making and Health Technology Assessment, Department of Public Health, Health Services Research and Health Technology Assessment, UMIT - University for Health Sciences, Medical Informatics and Technology, Eduard Wallnoefer Center I, A-6060 Hall i.T., Austria, Tel: +43 (0) 50-8648-3930, Fax: +43 (0) 50-8648-673931, e-mail: public-health@umit.at

Background: Biomarker tests are increasingly being offered by laboratories and clinicians as self-pay health services to screen asymptomatic individuals; however, sufficient evidence may not be available to support this practice. We investigated the benefit-harm tradeoffs associated with 11 biomarkers currently offered in Germany as self-pay tests to screen for cancer. Methods: We systematically searched bibliographic databases for health technology assessments, systematic reviews and randomized-controlled trials (RCTs) through September 2015. We included publications that analysed cancer screening biomarkers and reported patient-relevant outcomes (mortality, morbidity, quality of life), and potential harms of screening, among asymptomatic individuals in screening and non-screening arms. Language was restricted to English and German. Two reviewers independently screened references; data were extracted and quality of included studies was evaluated by a reviewer and validated by a second reviewer. Results: Six publications of secondary literature and four publications reporting results from two RCTs were included. For 10 cancer screening biomarkers, no direct evidence on patient-relevant outcomes was available. Only one trial, which simultaneously assessed cancer antigen 125 (CA125) and vaginal ultrasound for ovarian cancer screening, provided the outcome of interest. Screening compared with usual care did not reduce ovarian cancer mortality. Patient harms included overdiagnosis and false-positive results. Conclusion: Although ovarian cancer screening with CA125 showed no benefit, false-positive tests, overdiagnosis and overtreatment were reported. Physicians and laboratories should provide patients with comprehensive information about the lack of evidence and potential harms caused by biomarker screening tests offered as a self-pay health service.

Background

Benefit catalogs that are defined by the decision-making bodies of the statutory health insurance or national health services vary considerably between European countries. Differences might occur through diverse judgement on medical need, varying procedures and methods in assessment and appraisal of the clinical effectiveness of treatments, but also due to resource constraints. Besides covered health-care services, insured persons gain access to further medical services through an additional insurance or self-pay agreement.

In Germany, individual health services (IGeL) are examinations and therapies that are not covered by the statutory health insurance (German: Gesetzliche Krankenversicherung), requiring patients to pay out-of-pocket, if utilized. Data on the frequency and quality of these offers are not routinely collected and therefore remain mainly unknown. Studies report varying demands; up to 19% of insured persons asked for them, whereas between 19 and 53% have received offers without asking. Up to 80% of these offers were finally performed. The services that are most commonly used include intraocular pressure, which represents up to 40% of services, and ultrasound scans with up to 25%, followed by cancer screening and blood or laboratory tests with 9.7–14% of IGeL.

These services are often considered opportunistic screening because they are predominately offered to healthy, asymptomatic individuals outside an established screening program. A women’s health survey of physicians in the USA revealed that 6.3% [95% confidence interval (CI): 4.4–8.9] would routinely order or offer ovarian cancer screening for low-risk women and 24.0% (CI: 20.5–28.0) for medium-risk women.

Screening tests are used to define an individual’s risk of a particular disease or condition at an early stage for otherwise healthy individuals. An abnormal screen result can be followed up by additional testing to obtain a diagnosis. The primary aim of cancer screening is to reduce cancer-specific mortality and to detect and treat cancer at an earlier stage in order to improve the patient’s prognosis. Only the individuals with a true positive diagnosis following an abnormal result have the opportunity to benefit from screening. However, a true positive test result does not eliminate potential screen-related harms, which can include overdiagnosis (i.e. an individual would not have been clinically affected by their cancer in their lifetime). The prevalence of a disease in the general
population will impact the proportion of individuals that could potentially benefit from screening, but also the proportion that may be harmed as a result of being screened.\(^5,10\) Consequently, already in 1968, Wilson and Jungner established ‘principles and practice of screening for disease’, which defined conditions that should be met prior to implementing a candidate screening program.\(^7\) These screening criteria were adapted by the UK National Screening Committee and cover a broad range of screening-related factors including: the condition, the test, and the treatment and the program itself.\(^11\) Of utmost importance is the principle that the benefit of screening should outweigh the potential harm. Ideally, evidence to warrant a population-based screening program should be documented in randomized-controlled trials (RCT) of high methodological quality.\(^6,8,11\)

In Germany, the evaluation of efficacy and effectiveness of screening is regulated by the Federal Joint Committee (G-BA). According to the Sections 25 und 26 SGB V (Social Code), the G-BA decides which services will be reimbursed through statutory health insurance. Currently, screening of predefined risk-groups covered by the statutory health insurance include breast, cervical, skin, prostate and colorectal cancer.\(^12\) Such screening programs are funded and incorporate quality control measures and surveillance.\(^13\) Furthermore, professional societies establish cancer-specific clinical practice guidelines that provide recommendations on how to screen, treat and diagnose each cancer.\(^14\)

In contrast, screening with tests (e.g. biomarker tests) not covered by statutory health insurance (and require self-pay) is not guided by any predefined evidence-based screening criteria or even standardized recommendations because quality-assurance for IGel is not required.\(^4\) Therefore, biomarker tests, offered as a self-pay service do not have to provide documented effectiveness prior to being offered to patients.\(^4\)

The ‘objective’ of this study was to investigate the clinical effectiveness of biomarkers (table 1) that are currently offered as a self-pay healthcare service for cancer screening of asymptomatic individuals. The outcomes of interest were the effect of screening on cancer-specific mortality and morbidity, quality of life and harms, e.g. the screening procedure, false-negative test results as well as (false and true) positive results including follow-up diagnostics and examinations leading to overdiagnosis and overtreatment.

### Methods

#### Biomarker selection

Candidate screening biomarkers were identified through a web-based search exploring physician- and laboratory-homepages in Germany. Three search steps were performed (figure 1). In each step combinations of different search terms in German language were used. Finally, we included cancer screening biomarkers that were offered as a self-pay service at least eight times within the first 200 references identified through the Google search engine (table 1). Frequently offered biomarker tests that have already been evaluated for clinical effectiveness were excluded [i.e. prostate-specific antigen\(^15\) and human papillomavirus test\(^16\)].

#### Literature search

The systematic literature search involved a two-step process. First, health technology assessments (HTA) and systematic reviews (SRs) were searched for each identified biomarker test using MEDLINE, EMBASE, The Database of Abstracts of Reviews of Effects, the databases of the German Agency for Health Technology Assessment and of the International Network of Agencies for Health Technology Assessment. Second, we searched for RCTs using MEDLINE, EMBASE and The Cochrane Library. In order to supplement our search, we reviewed included references to identify additional sources. Keywords and Medical Subject Headings for biomarker, cancer, screening and publication type as well as population and study type for RCTs were used to develop the search algorithm (Supplementary Appendix table S3). Literature management was performed using ENDNOTE X5.

#### Inclusion criteria and literature selection

HTA reports and SRs were included if one or more biomarkers were evaluated as cancer screening test in an asymptomatic control population treated with usual care and compared with a screened asymptomatic population with the biomarker test of interest. Outcomes of interest included cancer-specific mortality and morbidity, quality of life and the potential harms of screening. Language was restricted to English or German. To identify relevant RCTs, consistent inclusion criteria to that of the previous

### Table 1 Analysed biomarkers offered as self-pay service for cancer screening and clinical practice guideline recommendations by cancer location

<table>
<thead>
<tr>
<th>Biomarker and Cancer location</th>
<th>Breast</th>
<th>Ovary</th>
<th>Testis</th>
<th>Prostate</th>
<th>Bladder</th>
<th>Gut</th>
<th>Liver</th>
<th>Lung</th>
<th>Pancreas</th>
</tr>
</thead>
<tbody>
<tr>
<td>AFP</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA125</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA15-3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CA19-9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CEA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CYFRA21-1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hHCG</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NMP22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>M2-PK</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NSE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PCA3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Guideline recommendation* CEA/CA15-3 not mentioned at all\(^6\) | No guideline available | X | PCA3 not mentioned at all\(^9\) | No guideline available | CEA/CA19-9 no screening discussed\(^25\) | No screening discussed\(^31\) | All NR\(^25\) | CEA no screening discussed\(^32\) | CA19-9 NR\(^32\)

RCT for screening none 2 RCT\(^25,26\) none none none none none none none none

---

\(^a\) German S3/S2- clinical practice guidelines of professional societies.
search were applied; however, the time period was restricted to identify only the most recent RCTs not included in existing HTAs or SRs of good quality.

The literature search was performed in an iterative manner, that means, literature was first selected by titles and abstracts, and only relevant references were screened in full text. The screening of references was carried out by two independent reviewers. Reasons for exclusion of full texts were documented (Supplementary Appendix table S4). No literature was excluded due to language restriction.

Data synthesis and quality assessment
Data synthesis was compiled into structured evidence tables. Two reviewers independently evaluated the methodological quality of the HTAs and SRs using the ‘assessment of multiple systematic reviews’ tool for HTAs and SRs. This table was modified through the addition of subdivided questions in order to provide a more detailed overview. For the RCTs, we evaluated the risk of bias using a standardized tool provided in ‘The Guidelines Manual 2009’ of the National Institute for Health and Clinical Excellence. A single reviewer extracted the data, which was verified for accuracy by a second reviewer.

Results
Biomarker selection
Out of 23 potential biomarkers that were offered for cancer screening and have not already been frequently evaluated for clinical effectiveness, 11 biomarkers were eight times or more often offered on German physician- and laboratory-homepages. These biomarkers were included in our analysis (table 1).

Literature search results
After removal of duplicates, the HTA and SR literature search yielded 2075 references, which were screened using the titles and abstracts. Of the 64 references identified for full text review, seven were not accessible. A total of five articles met our inclusion criteria, i.e. one HTA, two SRs, and two evidence-update reports from the US Preventive Services Task Force. All five publications investigated the biomarker cancer antigen 125 (CA125) as a screening test for ovarian cancer. One evidence synthesis report, evaluating bladder cancer screening using the biomarker nuclear matrix protein 22 (NMP22), was also included. For the remaining nine biomarkers, no secondary studies that met the inclusion criteria were identified.

The objective of all five included publications investigating ovarian cancer screening was to synthesize the available evidence on the benefits and harms of screening examinations using transvaginal ultrasound (VUS) and/or CA125 measurement. The HTA conducted by Bell et al. included studies performed through 1996, the review by Myers et al. included studies conducted between 1966 and 2006, while the review by Reade et al. searched studies until 2012, respectively. All three publications were of good methodological and reporting quality (Supplementary Appendix table S5). Although the review of Reade and colleagues provided evidence on overall ovarian cancer screening, the provided results were not clearly distinct in terms of screening strategy, results and outcomes for our research purposes. The last two reports we identified were self-defined as brief evidence updates, therefore, the methods as well as results were only shortly described.

The report on bladder cancer investigated whether there is direct evidence that supports the effect of bladder cancer screening on reducing cancer mortality or morbidity. The report utilized several bibliographic databases searched until year 2009/10, and covered a broad spectrum of screening methods, including urine cytology and urine biomarkers such as NMP22. The methodological quality of the report was sound (Supplementary Appendix table S5).
Based on the HTA/SR results, we targeted the search to identify new CA125 RCTs conducted after 2005 until September 2015. Due to the limited number of references identified in our search for NMP22 RCTs ($n = 38$), all these references were screened through September 2015. For the remaining nine biomarkers, we searched the bibliographic databases for RCTs conducted through September 2015. Following full text review, only four references met inclusion criteria; all identified studies evaluated the ovarian cancer biomarker CA125 (figure 2).

For 10 biomarkers, no direct evidence on patient-relevant outcomes investigated by RCTs was available, while two RCTs for the CA125 ovarian cancer screening biomarker were identified.\textsuperscript{25,26} Review of the included HTAs and SRs for NMP22, as well as for CA125, did not reveal any additional RCTs that could have been included. The primary conclusion, consistent across four of the five secondary studies, was that the evidence-base on patient-relevant outcomes such as mortality and morbidity is missing and it is recommended to conduct studies with a low risk of bias for example RCTs to assess benefits and harms of screening. The fifth review reported that there are no improvements in patient relevant outcomes through ovarian cancer screening with VUS and CA125, but harm occurs through false-positive results.\textsuperscript{21}

Figure 2 Flow chart of literature selection of randomized-controlled trials. AFP, alpha-fetoprotein; CA125, cancer antigen 125; CA15-3, cancer antigen 15-3; CA19-9, cancer antigen 19-9; CEA, carinoembryonic antigen; Cyfra21-1, cytokeratin 19 fragment antigen 21-1; β-HCG, β-subunit of human chorionic gonadotropin; NMP22, nuclear matrix protein 22; M2-PK, M2-pyruvate kinase; NSE, neuron-specific enolase; PCA3, prostate cancer antigen 3.
### Results of trials for ovarian cancer screening

The two multicenter ovarian cancer screening RCTs included were the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO),\(^{25,27}\) which recruited patients across the USA, and the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)\(^{26}\) performed in England, Wales and Northern Ireland. In both trials the primary outcome of interest was ovarian cancer-specific mortality. The screening strategy in the PLCO trial was VUS with concurrent CA125 measurement, and the UKCTOCS trial included two intervention arms involving different screening approaches. The first screening strategy involved VUS only (ultrasound screening strategy (USS) group), while the second screening strategy involved CA125 measurement (multimodal screening strategy (MMS)). Both strategies (MMS and VUS) had two levels differing by screening interval length (Supplementary Appendix table S6). In the MMS, the screening intervals of the level 1 were based on a predefined risk algorithm comprising of the woman’s age and CA125 measurement allowing classification of risk as normal, intermediate or elevated.\(^{26}\) Level 2 of the MMS screening approach was a vaginal ultrasound examination. The PLCO trial randomized 78,216 women aged 55–74 years into two groups, and the UKCTOCS trial randomized 202,638 postmenopausal women between 50 and 74 years into two intervention groups (USS or MMS) and a single control group. In both trials the control group received usual care. The protocol and conduct of both trials were of high quality, and the risk of bias in the PLCO trial was assessed as low (Supplementary Appendix table S7). End of trial results were documented and resulted in unnecessary surveillance of individuals as normal, intermediate or elevated.\(^{26}\) Level 2 of the MMS screening approach was a vaginal ultrasound examination. The PLCO trial randomized 78,216 women aged 55–74 years into two groups, and the UKCTOCS trial randomized 202,638 postmenopausal women between 50 and 74 years into two intervention groups (USS or MMS) and a single control group. In both trials the control group received usual care. The protocol and conduct of both trials were of high quality, and the risk of bias in the PLCO trial was assessed as low (Supplementary Appendix table S7). End of trial results were documented and resulted in unnecessary surveillance of individuals as normal, intermediate or elevated.\(^{26}\) Level 2 of the MMS screening approach was a vaginal ultrasound examination.

### Discussion

Primary evidence, provided by RCTs, was only found for one biomarker. Two RCTs analysed ovarian cancer screening with CA125. The PLCO trial showed no benefit on patient-relevant outcomes for asymptomatic women screened for ovarian cancer compared with an unscreened control group.\(^{25}\) In contrast, harms were documented and resulted in unnecessary surveillance of individuals with false-positive screening tests, overdiagnosis and overtreatment. Both RCTs provide evidence with a low risk of bias by enrolling a large number of participants, allowing a sufficient length of follow up, as well as, including a control group. Evidence gained through RCTs is considered of high importance in research for ovarian cancer screening as well as for other cancer types,\(^{11}\) but as our results revealed, RCTs for screening purposes with biomarkers are uncommon. Research is ongoing to improve and develop efficient screening strategies. Definition of high risk groups as well as optimizing the screening method and interval are key elements of improving screening.\(^{33,34}\) Over all, the evidence on mortality and morbidity of new strategies must be established in studies with a low risk of bias until screening can be provided to defined populations which may derive the most benefit.\(^{6,34}\)

Our review also revealed that, to our knowledge, there are no published RCTs assessing the effectiveness of 10 identified biomarkers that are currently available as a self-pay health service. For carcinoembryonic antigen (CEA), cytokeratin 19 fragment antigen 21-1 (CYFRA21-1) and neuron-specific enolase (NSE) biomarkers no evidence was found that investigated these markers as a screening test. There was no literature on the biomarkers β-subunit of human chorionic gonadotropin (β-HCG) and CA15-3 available which investigated cancer. We examined German clinical practice guidelines that provide evidence-based cancer-specific recommendations for cancer treatment as well as screening, for all indications of these biomarkers (table 1) and did not find a practice guideline that

### Table 2 Mortality, morbidity and diagnostic results (recent available results)

<table>
<thead>
<tr>
<th>Group</th>
<th>PLCO(^{25})</th>
<th>UKCTOCS(^{26})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Intervention (CA125+VUS)</td>
<td>Control</td>
</tr>
<tr>
<td>n analysed</td>
<td>34 253</td>
<td>34 304</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>1.4–1.8 (90–T5)</td>
<td>–</td>
</tr>
<tr>
<td>stage I</td>
<td>212 (5.7)</td>
<td>176 (4.7)</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>1.21 (CI:0.99–1.48)</td>
<td>–</td>
</tr>
<tr>
<td>stage II</td>
<td>32 (15)</td>
<td>18 (10)</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>1.01 (CI:0.96–1.06)</td>
<td>–</td>
</tr>
<tr>
<td>stage III</td>
<td>15 (7)</td>
<td>20 (11)</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>0.99 (CI:0.82–1.17)</td>
<td>–</td>
</tr>
<tr>
<td>stage IV</td>
<td>120 (57)</td>
<td>83 (47)</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>1.8 (CI:1.5–2.2)</td>
<td>–</td>
</tr>
<tr>
<td>No of deaths due to ovarian cancer (per 10 000 person-years)</td>
<td>43 (20)</td>
<td>54 (21)</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>1.0 (CI:0.8–1.2)</td>
<td>–</td>
</tr>
<tr>
<td>n (% woman with complications (including infections, complications during operation)</td>
<td>2924 (76.6)</td>
<td>2914 (76.2)</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>1.01 (CI:0.96–1.06)</td>
<td>–</td>
</tr>
<tr>
<td>No of deaths due to overall mortality (excluding ovarian, colorectal and lung cancer (per 10 000 person-years)</td>
<td>118 (3.1)</td>
<td>100 (2.6)</td>
</tr>
<tr>
<td>Rate ratio</td>
<td>0.99 (CI:0.82–1.17)</td>
<td>–</td>
</tr>
<tr>
<td>No of surgeries after false-positive result</td>
<td>1080</td>
<td>–</td>
</tr>
<tr>
<td>n (% woman with complications (including infections, complications during operation)</td>
<td>183 (15)</td>
<td>–</td>
</tr>
</tbody>
</table>

CA125, cancer antigen 125; CI, 95% confidence interval; MMS, multimodal screening; PLCO, Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial; T, screening rounds; UKCTOCS, United Kingdom Collaborative Trial of Ovarian Cancer Screening (Trial ongoing); VUS, transvaginal ultrasound.
recommended a single one of the investigated biomarkers for cancer screening in a general, asymptomatic population. For testical cancer (biomarker: β-HCG) as well as for bladder cancer [biomarker: NMP22], practice guidelines are not available. For example CEA, CYFRA21-1 andNSE are not recommended for screening purposes in the German clinical practice guideline for lung cancer.3,7 Similarly, CA15-3 and CEA were not mentioned in the German clinical practice guideline for breast cancer.36 Furthermore, in the clinical practice guideline of the American Society of Clinical Oncology (ASCO) CA15-3 is not recommended for screening, diagnosis or staging.37 CEA for gut or breast cancer is also not recommended by ASCO.37,38 For other biomarkers such as alpha-fetoprotein (AFP), M2-pyruvate kinase (M2-PK) and NMP22 SRs or primary studies are available for cancer screening, but this literature is restricted to high-risk populations, symptomatic populations, or did not provide any control group.24,39,40 Conclusions for screening of high-risk persons, particularly for patient-relevant outcomes, are not transferable to other lower-risk populations.

In Germany, there is a quality-assured screening program established for mammography and an evolving one for colorectal cancer. For these programs, great effort has been made to establish high-quality in order to reduce cancer-specific mortality while simultaneously minimizing the potential risk of screening.12,13 The statutory health insurance covers up to two colonoscopies for all persons above 55 years and the participation in mammography screening every 2 years for women aged 50–69 years.12 Even though mammography is a quality-assured program, it is still a widely debated screening program with opposing positions regarding the benefits and harms of this program.41,42 Despite the fact that these quality-assured screening programs are implemented, they are still debated. Therefore, offering biomarkers for breast cancer (CEA and CA15-3), as well as CA19-9, CEA and M2-PK for colorectal cancer screening, should be questioned. Even though German clinical guidelines do not recommend any of the investigated biomarkers, they continue to be offered in clinical practice, e.g. self-pay services recommended by physicians or laboratories to potential users. Claims data of biomarker usage for screening purposes as self-pay service are not routinely collected and it is at the physician’s discretion to choose which tests are offered.4 In a health survey on ovarian cancer screening with VUS and CA125 in the USA, 28.5% (CI: 24.5–32.9) reported not to adhere to current recommendations for women at low risk.5

Offering these services necessitates the provision of evidence-based information (or information about the lack of evidence) to the patient. Patients must be comprehensively informed about the potential benefit as well as the harms that may be a result of undergoing the test procedure, follow-up diagnostics or repeated tests.6 This information should guide the patient to make a decision for or against these services. Nevertheless, IGeL-tests cannot be compared with the screening strategies in RCTs. They are offered as single tests without a pre-defined screening interval and it is up to the physician to decide how to conduct diagnostics following receipt of the test result.

The Medical Advisory Service of Health Insurance [Medizinischer Dienst des Spitzenverbandes Bund der Krankenkassen e.V. (MDS)] introduced a service called ‘IGeL-Monitor’. The IGeL-Monitor provides information about IGeL-services in general as well as evidence-based reports about elective services (www.igele-monitor.de). For example, information about NMP22 as bladder cancer screening test as well as M2-PK test for colon cancer are available. Within these reports the benefit-harm balance is stated as ‘tendency to negative’ for NMP2244 and ‘unclear’ for M2-PK.45

Our analysis shows that for 10 biomarkers there is no sufficient evidence to support offering these biomarkers for screening purposes. Importantly, for CA125, screen-related harms were documented. Strategies to avoid such services without providing sufficient information should be established, in order to ascertain the possible harms as well as to prevent costs associated with unnecessary follow-up diagnostics, which other than the test itself, will be covered by the statutory health insurance. There should be ongoing discussions about whether or not to provide essential information to potential consumers of screening tests with large uncertainty regarding the benefit-harm balance. Such information can help protect consumers and health insurers from negative consequences.

Limitations

Our study has several limitations. The selection of investigated biomarkers used for cancer screening may not be representative for all biomarkers offered as self-pay services by physicians and laboratories for cancer screening. However, a person interested in biomarker cancer screening offered as self-pay service possibly would have found the investigated biomarkers identified via the Google search engine. Representative frequencies of using these biomarkers are not available, but it can be assumed that these biomarkers are offered regularly. Therefore, it was our aim to investigate the evidence on which these offers are based. Further, the literature search did not include searching for trials at clinicaltrials.gov. Conclusions about the benefits and harms for 10 cancer screening biomarkers could not be drawn due to the lack of RCTs conducted on these biomarkers. The trials that were available, but did not meet our inclusion criteria and were therefore excluded, did not investigate patient-relevant outcomes such as mortality or morbidity, did not provide a control group or reference test, or the patients were outside our target population, i.e. symptomatic or at high risk. Subsequently, these trials could not provide insight to answer our specific research question.

The PLCO trial did not analyse the outcomes of interest for CA125 without a combination with other tests; subsequently, the conclusions for screening with this biomarker test alone are limited. In addition, we are awaiting end of study results for the UKCTOCS trial. This trial applied a complex screening strategy using a predefined risk algorithm influencing the screening interval. It is uncertain whether or not this strategy is transferable to the self-pay screening context.

Conclusion

RCTs documenting the benefit-harm tradeoffs for 10 biomarkers that are currently offered outside statutory health insurance, are lacking. For ovarian cancer screening, CA125 yielded no benefit in terms of reducing cancer mortality, and also produced measurable screening-related harms. If providers decide to offer such biomarker tests, information to patients about the lack of evidence and potential harms should be explicit.

Supplementary data

Supplementary data are available at EURPUB online.

Acknowledgements

The authors would like to thank Theresa Hunger for the knowledge support regarding IGeL services.

Funding

This work was supported by the COMET Center Oncotyrol, which is funded by the Austrian Federal Ministries BMVIT/BMWJ (via FFG) and the Tyrolean Future Foundation/Standortagentur Tirol (SAT). The funding agreement ensured the authors’ independence in designing studies, data interpretation, writing and publishing results.
Conflicts of interests: All authors are or have been partly employed by Oncotyrol.

Key points
- The clinical effectiveness of self-pay biomarker screening tests is rarely investigated using rigorous randomized controlled trials.
- The biomarker tests currently offered in Germany as a self-pay screening test do not have documented proof of benefit or harms for the patients.
- Our review underscores the necessity of a critical review of biomarker screening tests offered as a self-pay health service. Providers of these tests should present patients with information about the evidence base when offering these tests.

References
Introduction

At least one in 1000 newborns are affected by hearing impairment,1–3 and the studies of hearing-impaired toddlers highlight the benefits of early diagnosis and intervention.4,5 Early intervention enhances language and speech outcomes as well as cognition and socioemotional development.6–8

Universal newborn hearing screening is recommended by international committees2,7,8, and recommendations from the Joint Committee on Infant Hearing include screening before 1 month of age, establishing a diagnosis before 3 months, and initiating interventions as soon as possible after diagnosis (no later than 6 months).7 A universal newborn hearing screening programme is essential to meet these goals. In practice, although the shared purpose is to diagnose hearing-impaired children early, different organisational designs could be implemented (e.g. the kind of tests, number of screening steps before diagnostic referral, newborns’ age when tests are performed, the professionals that...