Abstract: Laboratory tests are important tools in primary care, but their use is sometimes inappropriate. The aim of this review is to give an overview of interventions applied in primary care to optimize the use of laboratory tests. A search for studies was made in the MEDLINE and EMBASE databases. We also extracted studies from two previous reviews published in 2015. Studies were included if they described application of an intervention aiming to optimize the use of laboratory tests. We also evaluated the overall risk of bias of the studies. We included 24 studies. The interventions were categorized as: education, feedback reports and computerized physician order entry (CPOE) strategies. Most of the studies were classified as medium or high risk of bias while only three studies were evaluated as low risk of bias. The majority of the studies aimed at reducing the number of tests, while four studies investigated interventions aiming to increase the use of specific tests. Despite the studies being heterogeneous, we made results comparable by transforming the results into weighted relative changes in number of tests when necessary. Education changed the number of tests consistently, and these results were supported by the low risk of bias of the papers. Feedback reports have mainly been applied in combination with education, while when used alone the effect was minimal. The use of CPOE strategies seem to produce a marked change in the number of test requests, however the studies were of medium or high risk of bias.

Introduction

Laboratory tests are essential for screening, diagnosis and monitoring of diseases and are extensively used both in hospitals and primary care. It is generally recognized that the number of laboratory tests are increasing far more than the number of patients. The increased use of tests probably have benefited some patients through earlier diagnosis and treatment, however, their contribution to the quality of care also is debatable. This is especially the case when the usage is not based on clinical need and scientific evidence, but rather on habits, defensive medicine or similar causes [1].

Furthermore, the disproportional growth in laboratory utilization is considered economically unsustainable [2] and therefore it is important that the right tests are used in the right situation for the right patients. Thus, for all of these reasons a variety of interventional strategies have been investigated with respect to their capability to optimize the use of laboratory tests.

Primary care physicians are first in line to meet patients in the healthcare system, and they are responsible for a considerable proportion of the biochemical tests requested. Nevertheless, this sector is less well studied compared to the hospital setting regarding strategies to improve the application of laboratory tests in clinical situations.

To our knowledge, there are only two reviews [3, 4], both published in 2015, which have evaluated different interventions in terms of improving the use of laboratory tests in the primary care sector. Surprisingly, despite covering the same period, the two reviews have only four references in common of a total of seventeen.

In this review, we give an updated overview of studies of different interventions aiming to optimize clinical use of laboratory tests by primary care physicians, taking the
quality of the published studies into account. We summarize and compare the outcomes of different types of interventions with regards to the number of tests ordered, as well as changes in health care expenditures when reported by the authors.

**Methods**

**Search strategy**

The search was implemented in two electronic databases: Medline (1st January 2010–20th September 2019) and Embase (1st January 2010–20th September 2019). The collection period was limited from 2010 to the present due to the two “2015 reviews” [3, 4] considered to cover the previous period (1947–February 2014) and because this review is meant to give an overview of the interventions applied recently, since the use of the IT-requesting systems has supplanted the use of paper request forms used in the past. Nevertheless, studies reported in the two 2015 reviews which fulfill our inclusion criteria were also included and evaluated in the present review.

No limits on language, subject or type were placed on the database searches. The searches were designed to be comprehensive and to cover a broad range of the field, including the hospital setting, to ensure that all types of connections or collaborations with primary care appeared in the search. The search query consisted of terms considered important by the authors to describe the subject of the review. The search was tailored to the specific requirements of each database (Supplemental Material 1).

All citations were imported into the Endnote X9 citation manager with duplicate citations being removed automatically or manually when found later in the process. Citations were then imported into the web-based review software program Covidence [5] for subsequent title and abstract relevance screening and data characterization of the full articles.

All potential papers underwent a two-step screening process: (1) all citations were screened based on the title and abstract by two independent reviewers (SL, SA); (2) the full text of each citation was read and assessed for inclusion by the same two reviewers independently. Disagreements were solved by discussion.

**Criteria**

Studies were included in the review based on the following inclusion and exclusion criteria:

### Inclusion criteria

**Essential**

1. the setting in which the intervention(s) were applied was primary care.
2. one or more specific intervention(s) was applied in order to change the use of one or more laboratory tests;
3. effectiveness of the intervention(s) was reported by difference in number of produced test(s) with and without the intervention;

**Additional**

4. differences in test-related expenditure or healthcare costs with and without the intervention was reported.

### Exclusion criteria

1. no intervention was applied;
2. interventions focused on other care settings than primary care;
3. interventions were applied for other types of investigations than biochemical, immunological, genetic or microbiological laboratory tests;
4. quantitative results were not reported or could not be extracted from the paper.

**Data collection**

For each study reported in this review, we extracted information regarding the authors, country, year of publication, evaluation period, type of interventional strategy(ies) used, laboratory test(s) evaluated, study design and outcomes in terms of:

- number of tests requested before applying intervention(s);
- number of tests requested after applying intervention(s)
- difference in number of tests before and after intervention(s)

We evaluated the quality of the studies as the overall risk of bias translated to “high risk of bias”, “medium risk of bias” or “low risk of bias”, based on the Cochrane Effective Practice and Organization of Care (EPOC) suggested risk of bias [6] (Supplemental Material 2).

The study designs included: Randomized control trials (RCT) and cluster randomized controlled trials (CRCT); non-randomized control trials such as before and after controlled studies (CBA); before and after studies and prospective studies lacking a control group or control test (BA); and interrupted time series analysis (ITS) [7].
All the information was stored and handled in an Access database 2007–2010.

Calculations

The effectiveness of the interventions was reported as a relative number, corresponding to the effect on the group of tests following the application of the intervention.

Studies describing more than one intervention, were reported in this review with two or more results.

When the difference in tests used was reported as percentages, the study result was used directly (for example 20% being identical to 0.20).

Otherwise, if the results of a comparator (control group or test) was also reported in the study (before and after the intervention was applied), we adjusted the change in the intervention group with the change in the comparator before further calculations:

1. Calculating adjustment for intervention group before, based on changing in comparator:

   \[ \Delta \text{Comparator} = \text{comparator after} - \text{comparator before} \]

   Adjusted intervention group before

   \[ = \text{Intervention group before} + \Delta \text{comparator} \]

2. Calculating relative change in intervention group adjusted by the change in the comparator:

   \[ \frac{(\text{Intervention group after} - \text{Adjusted intervention group before})}{\text{Adjusted intervention group before}} \]

In case of the intervention applied on more tests in the same study, a weighted relative change was calculated based on the absolute number of the individual tests both for the comparator (before and after) and the intervention group (before and after). The weight of each tests was simply the number of the individual test:

\[ \text{Weighted result} = \sum \left[ \frac{\text{changes (adjusted) in intervention group} \times \text{number of test intervention group (after)}}{\text{total number of tests in intervention group (after)}} \right] \]

Types of intervention

The interventions used in the included studies were categorized as following:

Educational strategies (EDU)

*Education:* traditional education such as lectures, production and distribution of guidelines, audit, or guidelines integrated within clinical decision support systems (CDSS).

*Cost display:* cost of the test(s) is displayed in the requisition system or otherwise made clear to the requester during the requesting.

Feedback reports

Preparation of individualized reports showing the numbers of specific tests requested and/or graphic displays, most often compared with other requesters and/or previous time periods.

Computerized physician order entry (CPOE) strategies

Administrative changes: changes to the test requisition forms in the IT-System, such as changing order forms, making new order profiles or unbundling panels.

Soft-blocking: IT-based strategies that provide an interruptive “pop-up” alert that stop the workflow by requiring the user to acknowledge the alert or to cancel the ordering of the test.

Hard-blocking: a gate-keeping intervention that is used when a test is considered inadequate or repeated without clinical justifications and is therefore not processed – for example blocks caused by minimal retesting intervals.

A special type of hard-blocking is reflex testing where laboratory diagnostic algorithms automatically add orders of specific tests to a request or alternatively blocks the following test depending on the result of the initial test.

Results

Study selection

The electronic searches produced 9,985 studies (4,126 in Medline and 5,859 in Embase). After removing duplicates, 9,287 were screened at the title and abstract level. Four hundred and twenty-five articles were selected for full-text screening and of these, 332 studies were eliminated as they were not reporting original quantitative results relevant for the review topic. Eighty studies were excluded because the interventions were not applied to primary care.
The electronic search ended with 13 studies that met our pre-specified criteria and were included in the review. Seventeen studies were extracted from the two 2015 reviews [3, 4] and of those, six studies were excluded: three due to a different focus area, two because they were also included in our own search and one because no clear results were reported.

The selection process left a final pool of 24 studies (Figure 1) that were published between 1992 and September 2019.

**Study characteristics**

Table 1 summarizes the characteristics of the included studies. Five studies implemented more than one strategy to reduce the number of tests either to different pools of tests [8, 17, 26] or to the same pool of tests [12, 19] and therefore, the results of those studies have been split into two or three parts, respectively. Two studies used the same interventions in two different laboratory areas, and the results were therefore also reported in two sections [11, 28].

Table 2 includes studies that reported interventions applied for increasing the use of tests. One study investigated exclusively interventions aiming at increasing the use of (under-requested) tests [31] while three studies [8, 13, 24] reported, as a part of their investigation, both results related to over-use of tests as well as to under-used tests and therefore those results have been described separately in Table 2.

The studies included in this review were conducted in UK [12, 14, 22], Italy [27, 28, 30], The Netherlands [9–11, 20], Canada [26], USA [16], Israel [21, 23, 24], Sweden [13], Portugal [8], Spain [17–19, 29, 31], Finland [25], and New Zealand [15]. Although most were published in English, one was in Italian and another in Dutch.

The observational period in the 24 studies varied from three months to five years.

**Study design and risk of bias**

For 23 of the studies, the risk of bias was determined by following the criteria for RCT and non-RCT [6]. Seven studies had a high overall risk of bias, of which five were BA [21, 22, 24, 26, 31], and three CBA [15, 23, 27]. Thirteen were classified as having a medium risk of bias, of which 11 were CBA [9, 13, 17–20, 25, 28–31], and two were RCT [10, 14]. Only three studies were classified as having a low risk of bias, one was a CRCT [12] and two were RCT [8, 11].

To determine the quality of the Interrupted time series (ITS) study, the ITS studies criteria was followed [6], and the study [16] was evaluated as having a medium risk of bias.

Of the five RCT/CRCT studies, randomization was performed at the clinics level in three of the studies [10, 12, 14], at the server level in one [8] (which means that all physicians using the same server received the same intervention) and at the primary care physicians’ level in the last one [11].

Of the 12 non-RCT studies, five used a test as control [17, 18, 23, 25, 29], and seven used a control group [9, 13, 15, 24, 26, 28, 31].
<table>
<thead>
<tr>
<th>References</th>
<th>Author-country-year</th>
<th>Design</th>
<th>Intervention designed in collaboration with requestor</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Duration period</th>
<th>Tests</th>
<th>Cost evaluation</th>
<th>Relative reduction of targeted tests</th>
<th>Overall risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[8]</td>
<td>Martins et al., Portugal, 2017</td>
<td>RCT</td>
<td>No</td>
<td>CDSS</td>
<td>Control group</td>
<td>Six months pre-intervention, seven months intervention and eight months follow up</td>
<td>CA 19-9; cholesterol; glucose; hepatitis B surface antigen; Hepatitis C antibody; VDRL</td>
<td>Not reported</td>
<td>−0.02</td>
<td>Low</td>
</tr>
<tr>
<td>[9]</td>
<td>Beijer et al., The Netherlands, 2016</td>
<td>CBA</td>
<td>Yes</td>
<td>Audit</td>
<td>Control group</td>
<td>Six months pre-intervention and six months intervention</td>
<td>Ferritin; folate; hemoglobin; iron; LDH; leukocytes; reticulocyte; thrombocytes; transferrin; vitamin B12</td>
<td>Not reported</td>
<td>−0.09</td>
<td>Medium</td>
</tr>
<tr>
<td>[10]</td>
<td>van Wijk et al., The Netherlands, 2001</td>
<td>RCT</td>
<td>No</td>
<td>CDSS integrated with guidelines</td>
<td>Control group</td>
<td>Three months pre-intervention and one year intervention</td>
<td>Laboratory tests</td>
<td>Not reported</td>
<td>−0.2</td>
<td>Medium</td>
</tr>
<tr>
<td>[11]</td>
<td>Verstappen et al., The Netherlands, 2003</td>
<td>RCT</td>
<td>No</td>
<td>Guidelines, audit and feedback report</td>
<td>Control group</td>
<td>Six months pre-intervention, six months intervention and six months follow up</td>
<td>BUN; AST; LDH; amylase; alkaline phosphatase; bilirubin;</td>
<td>Not reported</td>
<td>−0.12</td>
<td>Low</td>
</tr>
<tr>
<td>[11]</td>
<td>Verstappen et al., The Netherlands, 2003</td>
<td>RCT</td>
<td>No</td>
<td>Guidelines, audit and feedback report</td>
<td>Control group</td>
<td>Six months pre-intervention, six months intervention and six months follow up</td>
<td>IgE; leukocyte count</td>
<td>Not reported</td>
<td>−0.08</td>
<td>Low</td>
</tr>
<tr>
<td>[12]</td>
<td>Thomas et al., UK, 2006</td>
<td>CRCT</td>
<td>No</td>
<td>CDSS and feedback report</td>
<td>Control group</td>
<td>One year pre-intervention and one year intervention</td>
<td>Auto antibody screen; Ca-125; CEA; ferritin; FSH; Helicobacter pylori serum; IgE; TSH; vitamin B 12</td>
<td>Not reported</td>
<td>−0.31</td>
<td>Low</td>
</tr>
<tr>
<td>[12]</td>
<td>Thomas et al., UK, 2006</td>
<td>CRCT</td>
<td>No</td>
<td>feedback report</td>
<td>Control group</td>
<td>One year pre-intervention and one year intervention</td>
<td>Auto antibody screen; Ca-125; CEA; ferritin; FSH; Helicobacter pylori serum; IgE; TSH; vitamin B 12</td>
<td>Not reported</td>
<td>−0.15</td>
<td>Low</td>
</tr>
<tr>
<td>[12]</td>
<td>Thomas et al., UK, 2006</td>
<td>CRCT</td>
<td>No</td>
<td>CDSS</td>
<td>Control group</td>
<td>One year pre-intervention and one year intervention</td>
<td>Auto antibody screen; Ca-125; CEA; ferritin; FSH; Helicobacter pylori serum; IgE; TSH; vitamin B 12</td>
<td>Not reported</td>
<td>−0.14</td>
<td>Low</td>
</tr>
<tr>
<td>References</td>
<td>Author-country-year</td>
<td>Design</td>
<td>Intervention designed in collaboration with requestor</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Duration period</td>
<td>Tests</td>
<td>Cost evaluation</td>
<td>Relative reduction of targeted tests</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------</td>
<td>--------</td>
<td>------------------------------------------------------</td>
<td>--------------</td>
<td>-----------</td>
<td>----------------</td>
<td>-------</td>
<td>----------------</td>
<td>----------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>[13]</td>
<td>Larsson et al., Sweden, 1999</td>
<td>CBA</td>
<td>No</td>
<td>Education program</td>
<td>Control group</td>
<td>Eight months pre-intervention, three months intervention and six months follow up</td>
<td>ALT/AST; bilirubin/alkaline phosphatase; Cholesterol/HDL; Cholesterol/total number; sodium/potassium; FT3/TSH; FT4/TSH</td>
<td>SEK 400,000 saved for assays that were recommended to decrease</td>
<td>−0.34</td>
<td>Medium</td>
</tr>
<tr>
<td>[14]</td>
<td>Baker et al., UK, 2003</td>
<td>RCT</td>
<td>No</td>
<td>Guidelines and feedback report</td>
<td>Control group</td>
<td>One year intervention</td>
<td>Lipids; viscosity; thyroid function; rheumatoid factor; urine culture</td>
<td>Not reported</td>
<td>0</td>
<td>Medium</td>
</tr>
<tr>
<td>[15]</td>
<td>Tomlin et al., New Zealand, 2011</td>
<td>CBA</td>
<td>No</td>
<td>Guidelines and feedback report</td>
<td>Control group</td>
<td>Two years pre-intervention and two years intervention</td>
<td>CRP; ESR; TSH; FT4; FT3; Culture; Giarda and Cryptosporidium; Ova and parasites</td>
<td>Saving in net expenditure 25.1%</td>
<td>−0.212</td>
<td>High</td>
</tr>
<tr>
<td>[16]</td>
<td>Horn et al., USA, 2013</td>
<td>ITS</td>
<td>Yes</td>
<td>Cost displays</td>
<td>Control group</td>
<td>One year pre-intervention and six months intervention</td>
<td>25-OH Vitamin D; α fetoprotein; ALT; BMP; BUN; BNP; CBC with differential; Chlamydia/GC genital screening; Chlamydia/GC urine; creatinine; CMP (screen); electrolytes; ESR; ferritin; glucose; group a beta hemolytic streptococcus; HbA1C; iron binding profile; lipid profile; micro albumin (urine); Pap test; PTH; PSA; TSH; tissue transglutaminase; urine culture; urine analysis</td>
<td>The average cost saving was $4,545 per 1,000 visits per month.</td>
<td>−0.01</td>
<td>Medium</td>
</tr>
<tr>
<td>Administrative</td>
<td>Salinas et al., Spain, 2014</td>
<td>CBA</td>
<td>Yes</td>
<td>Removing tests from profile</td>
<td>Control test</td>
<td>Two years pre-intervention and two years intervention</td>
<td>AST; GGT; phosphate</td>
<td>€34,064 saved post-intervention (the savings were calculated using the cost of extra test reagents)</td>
<td>−0.51</td>
<td>Medium</td>
</tr>
<tr>
<td>[8]</td>
<td>Martins et al., Portugal, 2017</td>
<td>RCT</td>
<td>No</td>
<td>Changing the shortcut menu</td>
<td>Control group</td>
<td>Six months pre-intervention, seven months intervention and eight months follow up</td>
<td>ESR; SPE; uric acid;</td>
<td>Not reported</td>
<td>−0.53</td>
<td>Low</td>
</tr>
<tr>
<td>References</td>
<td>Author-country-year</td>
<td>Design</td>
<td>Intervention designed in collaboration with requestor</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Duration period</td>
<td>Tests</td>
<td>Cost evaluation</td>
<td>Relative reduction of targeted tests</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------</td>
<td>--------</td>
<td>-------------------------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>----------------</td>
<td>-------</td>
<td>----------------</td>
<td>---------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>[18]</td>
<td>Salinas et al., Spain, 2014</td>
<td>CBA</td>
<td>yes</td>
<td>Removing the test from the profiles</td>
<td>Control test</td>
<td>One year pre-intervention and one year intervention</td>
<td>Uric acid</td>
<td>€ 8,190 saved that resulted from the decrease in the request of UA in the post-intervention period as compared to the pre-intervention period.</td>
<td>−0.7</td>
<td>Medium</td>
</tr>
<tr>
<td>[19]</td>
<td>Pellicer et al., Spain, 2018</td>
<td>CBA</td>
<td>No</td>
<td>Changing profiles</td>
<td>Control group</td>
<td>Six years pre-intervention and six months intervention</td>
<td>TPOab; Igab</td>
<td>Not reported</td>
<td>−0.65</td>
<td>Medium</td>
</tr>
<tr>
<td>[20]</td>
<td>Zaat et al., The Netherlands, 1992</td>
<td>CBA</td>
<td>No</td>
<td>Reducing the number of tests in the form</td>
<td>Control group</td>
<td>One year From 178 tests to 15 tests</td>
<td>Not reported</td>
<td>−0.18</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td>[21]</td>
<td>Vardy et al., Israel, 2005</td>
<td>BA</td>
<td>Yes</td>
<td>Changing form</td>
<td>Time period</td>
<td>Four months pre-intervention and four months intervention</td>
<td>Laboratory routine tests</td>
<td>Not reported</td>
<td>−0.02</td>
<td>High</td>
</tr>
<tr>
<td>[22]</td>
<td>Bailey et al., UK, 2005</td>
<td>BA</td>
<td>No</td>
<td>Changing form</td>
<td>Time period</td>
<td>Five years pre-intervention</td>
<td>Calcium, LDH; rheumatoid factors; CRP</td>
<td>Not reported</td>
<td>−0.51</td>
<td>High</td>
</tr>
<tr>
<td>[23]</td>
<td>Kahan et al., Israel, 2009</td>
<td>CBA</td>
<td>No</td>
<td>Changing form</td>
<td>Control test</td>
<td>Five months pre-intervention, five months intervention</td>
<td>Vitamin B 12, folic acid, ferritin</td>
<td>Not reported</td>
<td>−0.4</td>
<td>High</td>
</tr>
<tr>
<td>[24]</td>
<td>Shalev et al., Israel, 2009</td>
<td>BA</td>
<td>No</td>
<td>Removing tests from the form</td>
<td>Time period</td>
<td>One year pre-intervention and two years intervention</td>
<td>TESTS DELETED: AST; GGT; total bilirubin; uric acid; Phosphorus; CK; PTT; CMV; EBV; antinuclear-ab; Hepatitis B-ab; feces; FSH; Beta-HCG quantity; LH; Hepatitis B core antibodies; E2; Prolactin; progesterone, testosterone; rubella antibodies IgG; cortisol; hepatitis a IgM; Sputum culture; puss culture; vagina culture; TESTS UN-CHANGED IN THE ORDER</td>
<td>The costs were marginally higher after the intervention period</td>
<td>−0.38</td>
<td>High</td>
</tr>
<tr>
<td>References</td>
<td>Author-country-year</td>
<td>Design</td>
<td>Intervention designed in collaboration with requestor</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Duration period</td>
<td>Tests</td>
<td>Cost evaluation</td>
<td>Relative reduction of targeted tests</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>-------------</td>
<td>----------------------</td>
<td>--------</td>
<td>--------------------------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-----------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------</td>
<td>---------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>FORM: ALT; creatinine; CBC; glucose; cholesterol; triglyceride; urea bun; HDL; LDL; alkaline phosphatase; Potassium; Sodium; urine general; TSH; Calcium; Urine culture; albumin; total protein; ESR; PT; HbA1c; throat culture; stool culture; beta-HCG quality</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Administrative + Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[25]</td>
<td>Seppanen, et al., Finland, 2016</td>
<td>CBA</td>
<td>No</td>
<td>Removing test from the order form + audit</td>
<td>Control test</td>
<td>Four years pre-intervention and four years intervention</td>
<td>AST; ESR</td>
<td>Not reported</td>
<td>–0.9 Medium</td>
<td></td>
</tr>
<tr>
<td>[26]</td>
<td>Elenenaei et al., Canada, 2016</td>
<td>BA</td>
<td>No</td>
<td>Unbundling panel + multifaced educational approach</td>
<td>Time period</td>
<td>Three months</td>
<td>Chloride, CO₂</td>
<td>$42,500 for chloride and $48,000 for CO₂ in annual marginal cost savings.</td>
<td>–0.72 High</td>
<td></td>
</tr>
<tr>
<td>[27]</td>
<td>Barìcchi et al., Italy, 2012</td>
<td>CBA</td>
<td>Yes</td>
<td>New profiles and audit</td>
<td>Control group</td>
<td>One year pre-intervention and one year intervention</td>
<td>Laboratory profiles: Normal adult profile; patient with myeloma; Monoclonal gammopathy; active chronic hepatitis; thyroid; hypertension; estrogen–progestogenic treatment</td>
<td>Not reported</td>
<td>–0.05 High</td>
<td></td>
</tr>
<tr>
<td><strong>Soft blocking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[28]</td>
<td>Camerotto et al., Italy, 2012</td>
<td>CBA</td>
<td>Yes</td>
<td>Interrupted pop-up alert</td>
<td>Control group</td>
<td>Six months intervention</td>
<td>110 laboratory tests</td>
<td>Not reported</td>
<td>–0.26 Medium</td>
<td></td>
</tr>
<tr>
<td>[28]</td>
<td>Camerotto et al., Italy, 2012</td>
<td>CBA</td>
<td>Yes</td>
<td>Interrupted pop-up alert</td>
<td>Control group</td>
<td>Six months intervention</td>
<td>110 laboratory tests</td>
<td>Not reported</td>
<td>–0.38 Medium</td>
<td></td>
</tr>
<tr>
<td>References</td>
<td>Author-country-year</td>
<td>Design</td>
<td>Intervention designed in collaboration with requestor</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Duration period</td>
<td>Tests</td>
<td>Cost evaluation</td>
<td>Relative reduction of targeted tests</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>------------</td>
<td>---------------------</td>
<td>--------</td>
<td>-----------------------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>-----------------</td>
<td>-------</td>
<td>----------------</td>
<td>----------------------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Hard blocking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[19]</td>
<td>Pellicer et al., Spain, 2018</td>
<td>CBA</td>
<td>No</td>
<td>Gate keeping</td>
<td>Control group</td>
<td>Six years pre-intervention and six months intervention</td>
<td>TPOab; tgab</td>
<td>Not reported</td>
<td>−0.13</td>
<td>Medium</td>
</tr>
<tr>
<td>[17]</td>
<td>Salinas et al., Spain, 2014</td>
<td>CBA</td>
<td>Yes</td>
<td>Reflexive test</td>
<td>Control test</td>
<td>Two years pre-intervention and two years intervention</td>
<td>IgA anti-gliadin antibody; iron; transferrin; total bilirubin</td>
<td>€34,064 saved post-intervention (the saving was calculated using the cost of extra tests reagents)</td>
<td>−0.15</td>
<td>Medium</td>
</tr>
<tr>
<td><strong>Hard blocking + education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[26]</td>
<td>Elnenaei et al., Canada, 2016</td>
<td>BA</td>
<td>No</td>
<td>Gatekeeping + multifaceted educational approach</td>
<td>Not reported</td>
<td>One year pre-intervention and one year intervention</td>
<td>25-OH Vitamin D; CEA; FOB; folate; LDH; SPE</td>
<td>$52,298 annual marginal cost savings</td>
<td>−0.51</td>
<td>High</td>
</tr>
<tr>
<td>[29]</td>
<td>Salinas et al., Spain, 2015</td>
<td>CBA</td>
<td>No</td>
<td>Gate keeping and comment display</td>
<td>Control test</td>
<td>Six years pre-intervention and four years intervention</td>
<td>1,25-(OH)2 vitamin D</td>
<td>The strategies resulted in saving €1,200 (no specified which costs)</td>
<td>−0.514</td>
<td>Medium</td>
</tr>
<tr>
<td>[30]</td>
<td>Caldarelli et al., Italy, 2017</td>
<td>CBA</td>
<td>No</td>
<td>Reflexive test + information letter to GPs</td>
<td>Time period</td>
<td>One year pre-intervention and two years intervention</td>
<td>TSH; FT4; FT3</td>
<td>€11,079 saved after intervention (calculated considering only the re-agents- cost of a single test was obtained by dividing the cost of the kit by the number of the test that can be done by 1 kit)</td>
<td>−0.10</td>
<td>Medium</td>
</tr>
<tr>
<td>Reference</td>
<td>Author-Country-year</td>
<td>Design</td>
<td>Intervention designed in collaboration with requestor</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Duration period</td>
<td>Tests</td>
<td>Cost evaluation</td>
<td>Relative increase of targeted tests</td>
<td>Overall risk of bias</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------</td>
<td>--------</td>
<td>-------------------------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>----------------</td>
<td>-------</td>
<td>-----------------</td>
<td>---------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[13] Larsson et al., Sweden, 1999</td>
<td>CBA</td>
<td>No</td>
<td>Educational program</td>
<td>Control group</td>
<td>Eight months pre-intervention, three months intervention and six months follow up</td>
<td>Calcium/total number; methylal/total number; ferritin/total number; TSH/total number; HbA1c/total number; triglycerides/cholesterol; U-albumin/total number</td>
<td>SEK 140,000 cost spent for assays that were recommended to increase</td>
<td>0.013</td>
<td>Medium</td>
<td></td>
</tr>
<tr>
<td><strong>Administrative</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[8] Martins et al., Portugal, 2017</td>
<td>RCT</td>
<td>No</td>
<td>Adding tests</td>
<td>Control group</td>
<td>Six months pre-intervention, seven months intervention and eight months follow up</td>
<td>FOB; HDL; triglycerides</td>
<td>Not reported</td>
<td>0.08</td>
<td>low</td>
<td></td>
</tr>
<tr>
<td>[24] Shalev, Israel, 2009</td>
<td>BA</td>
<td>No</td>
<td>Adding tests</td>
<td>Time period</td>
<td>One year pre-intervention and two years intervention</td>
<td>Occult blood; bilirubin (neonatal)</td>
<td>The costs were marginally higher after the reference period</td>
<td>0.94</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td><strong>Hard blocking</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[31] Salinas et al., Spain, 2019</td>
<td>BA</td>
<td>Yes</td>
<td>Reflexive test</td>
<td>No comparator</td>
<td>Six months</td>
<td>Albumin to creatinine ratio (ACR)</td>
<td>The costs of strip analysis was 275.8€ but the saving in albumin reagent was 1450.3€</td>
<td>0.079</td>
<td>High</td>
<td></td>
</tr>
</tbody>
</table>
19, 20, 27, 28]. Five studies used a time period as comparator [21, 22, 24, 26, 30], while one prospective study did not have neither control group/test nor control period [31]. The ITS study [16] used a control group as comparator.

**Intervention components**

The majority of the studies (66%) used a single intervention approach, while the remaining studies used a combination of interventions.

Results obtained by different types of strategies combined with the risk of bias are displayed graphically in Figures 2 and 3. It is clearly illustrated that the risk of bias is lower for studies investigating the educational strategies, showing relative changes of approximately 10–30% reduction in test numbers. The medium risk of bias studies were represented in all categories and showed results ranging from approximately +15% to −90%. The high risk of bias studies were mainly included in the administrative category and the range of results were very broad going from +94% to −72%.

Studies published before 2010 reported only educational strategies (half of them supplemented by feedback reports) or administrative changes, while studies published later than 2010 implemented more technically complicated interventions (Table 3).

**Impact of interventions aiming to decrease the number of tests**

**Educational strategies**

**Education**

The studies examining interventions based exclusively on an educational component were generally medium or low risk of bias, except from one study with a high risk of bias. Duration period of the applied interventions ranged from three months to two years.
Table 3: Distribution of interventions applied before and after 2010.

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Number of interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤2010</td>
<td></td>
</tr>
<tr>
<td>Administrative</td>
<td>6</td>
</tr>
<tr>
<td>Education</td>
<td>4</td>
</tr>
<tr>
<td>Education + feedback report</td>
<td>5</td>
</tr>
<tr>
<td>&gt;2010</td>
<td></td>
</tr>
<tr>
<td>Administrative</td>
<td>5</td>
</tr>
<tr>
<td>Administrative + education</td>
<td>3</td>
</tr>
<tr>
<td>Education</td>
<td>3</td>
</tr>
<tr>
<td>Education + feedback report</td>
<td>1</td>
</tr>
<tr>
<td>Hard blocking</td>
<td>3</td>
</tr>
<tr>
<td>Hard blocking + education</td>
<td>3</td>
</tr>
<tr>
<td>Soft blocking</td>
<td>2</td>
</tr>
</tbody>
</table>

Studies which applied traditional educational interventions such as audit or distributed guidelines reported results ranging from −0.08 to −0.34 [9, 11, 13, 15]. The overall conclusions of the audit and guidelines interventions, were generally positive, with the highest reduction in number of tests being achieved by applying an educational program, consisting of a two days lecture series with distributions of the references used during the lectures and a follow-up meeting six months after the lectures [13].

Three studies which integrated guidelines in the CDSS resulted in a reduction ranging from −0.02 to −0.31, where the two RCT low risk of bias studies achieved reductions of −0.02 and −0.31 [8, 12].

Economical outcome for educational interventions were evaluated by two studies. Tomlin et al. [15] reported the costs in terms of savings in net expenditures (up to 25%), while Larsson et al. [13] reported the direct laboratory costs saved (SEK 400.000).

**Costs displayed**

The study by Horn et al. [16] was the only one to examine the effect of costs display as a “stand alone” intervention, showing no effect on number of tests requested (−0.01).

**Feedback reports**

Feedback reports were predominantly combined with educational interventions [11, 15] and only two studies evaluated also the effect of the feedback reports by itself. Thomas et al. [12] assessed the feedback reports both alone (−0.15) and combined with guidelines (−0.31), while Baker et al. [14] found that feedback reports initially combined with distribution of guidelines did not have any influence on tests ordered by GPs.

**CPOE component**

**Administrative changes**

Nine studies applied administrative changes exclusively, including removing or changing tests in profiles or changing tests in the requisition form or shortcut menu [8, 17–24], while three studies combined these interventions with educational components [25–27].

The duration periods ranged from three months to five years.

The risk of bias for these studies were predominantly high [21–24] or medium [17–20], except one study with low risk of bias [8], and the results showed a very broad range (from −0.02 to −0.9).

Studies that removed tests from profiles were all classified as medium risk of bias achieving comparable reductions: (−0.51; −0.64 and −0.70, respectively) [17–19], while making modifications in the requisition order form resulted in widely different results, ranging from −0.02 to −0.53 [8, 20–24]. The quality of these papers were very heterogeneous as well.

In three studies, two with high risk of bias [26, 27] and one with medium risk of bias [25], the administrative changes were combined with multifaceted educational approaches (such as training sessions, distribution of guidelines, and reminders) [26] or audits [25, 27], achieving very different reductions, −0.72; −0.05; and −0.9, respectively.

Four studies reported the expenditures associated with the administrative approaches, described as either marginal costs of tests or by differences in costs for test reagents. Reports of marginal costs showed savings between 42,500 and 48,000$ in one study [26] and marginally higher costs after the intervention period in another [24]. Differences in costs for test reagents showed savings in both studies, with savings of 34,064€ and 8.190€, respectively [17, 18].

**Soft-blocking**

One medium risk of bias study used the soft blocking approach [28]. It consisted of a pop-up alert that proposed indications for the use of the included tests. GPs were free to override the alert, but with the obligation to justify the order. This approach was applied to two groups of GPs attached to two different hospitals laboratories. The strategy was considered successful as the number of requests reduced consistently in both groups compared to the control group (−0.26 and −0.38, respectively).

**Hard-blocking**

Six studies (four of medium risk of bias and two high risk of bias), used more rigid approaches either by using gate
keeping interventions that rejects tests from the clinicians orders [19, 26, 29] or by applying reflex testing algorithms [17, 30, 31].

In three of the studies, hard blocking was combined with education [26, 29, 30]. The studies that used gate keeping combined with multifaceted educational approaches (such as audit or letters which include educational components sent to the highest requesting GPs and memorandums) [26] or educational messages displayed to the GPs requisition system [29] achieved larger effects compared to gate keeping alone [19] (−0.51 and −0.51 vs. −0.13). In contrast, results obtained by reflex testing seemed to be independent of combination with educational interventions [30] or not [17] (−0.10 and −0.15, respectively).

In four studies, also costs information was reported, as either the cost of reagents [17, 29, 30] or as marginal costs of the test [26]. The reagent costs saved after reflex testing was applied, were 11,079€ [30] and 34,064€ [17] while, using a gate keeping approach, saved 1,200€ in reagents costs [29] and 52,298€ in marginal costs [26].

**Impact of interventions aiming to increase the number of tests**

Three studies were designed to apply interventions to increase the use of specific laboratory tests in addition to the strategies to decrease the use of other tests [8, 13, 24]. Only one study was designed to apply exclusively an intervention to increase the use of the evaluated test [31].

Results of interventions aiming to increase the number of tests are reported in Figure 3.

**Education**

One medium risk of bias study [13] applied an educational method to recommend the use of specific tests, however, the impact of the strategy achieved modest effect (0.08).

**CPOE component**

**Administrative changes**

Two studies examined also the effect of adding specific tests to the prescription format. They found that the number of tests increased modestly in the low risk of bias study [8] (0.08) while the increase achieved was very high (0.94) in the study, which had a high risk of bias [24].

**Algorithm (Hard-blocking)**

One high risk of bias study [31] evaluated the effect of adding albumin to the creatinine ratio (ACR) automatically in patients with hypertension if the test had not been requested in the previous year, according to the guideline. Authors reported that the algorithm automatically added +0.13 tests, which otherwise would have been missed.

**Studies designed together with requestors**

Eight studies designed the intervention in consensus with the requestors [9, 16–18, 21, 27, 28, 31] and they are represented in all the intervention categories. The span of the results were very broad, ranging from +0.07 to −0.51. Moreover, none of the studies had a follow-up period.

**Discussion**

We here present a systematic review of investigations of different interventions applied in primary care in order to optimize the use of laboratory tests. This review was designed to adjust the results of the interventions to make them comparable in each category. We summarized 24 studies, which reported 35 results of interventions, evaluating the effect of different strategies based on educational components, feedback reports or CPOE components or combinations of them to optimize test utilization.

We only found eight intervention results being evaluated as a high quality study due to a low risk of bias. While many studies are of lower quality being classified as medium or even high risk of bias, making results less valid.

**Sustainability/applicability**

High quality studies were primarily represented in papers evaluating educational components, and they often resulted in changes in test usage of 10–30%. The duration periods were however, generally short with most of the studies reporting interventions of less than one year, and most of the reports lacked a follow-up period, so it is unknown whether the changes are stable or the requesting patterns return to previous levels after the intervention period, with one outstanding exception: The long-term effect of the study of Larsson et al. [13] was evaluated after eight years [32] and showed that short continuation courses on optimal use of clinical chemistry can achieve permanent changes in the test ordering patterns of GPs.
One of the key roles of the laboratory is to provide an easy and direct access to guidelines and educational material for examining and treating the patients in the best possible way [33]. However, educational approaches often require a large amount of effort over time as well as dedicated staff and still may not be sufficient on their own [1, 34, 35].

We found only one study which applied a strategy where the cost of the tests are displayed in real time when the GP is ordering the tests [16], and it achieved insignificant results. In the hospital setting, this strategy has resulted in conflicting results, with some studies reporting that displaying the cost of the tests led to a decrease in the ordering of tests [36, 37] while other authors concluded the contrary [38]. Displaying costs on the GPs requisition form is still sparsely investigated, but GPs tend to use more common and cheaper tests compared to hospitals, which tend to request the majority of special and expensive tests. It is therefore likely that this intervention might have only minor effect on the requesting patterns of GPs.

Feedback reports on their own seem to have modest or no effect on requests made in primary care. This is in contrast to findings in the hospital setting, where the use of feedback reports (described in combination with educational components) have been found to have clear and prolonged effects on physician laboratory test-orders [39, 40]. However, as described by Pantoja et al. [41] the effect of feedback reports probably depend on a number of specific factors such as: How the reports are delivered to the GPs; whether they provide generic or specific information; whether they provide an explanation of their content or not; whether they require the healthcare professional to record a response or whether the targeted clinicians are involved in their development.

Many of the administrative interventions investigated alone or in combination with educational components were technically provided by the CPOE systems as changes of requesting profiles, new order profiles or unbundling panels.

Although administrative interventions achieved the greatest relative reductions, the range of the results differed widely, probably due to the heterogeneity of the studies and the high risk of bias of many of the studies, which makes the results uncertain. It was reported by the authors of the studies included in this review, that this kind of interventions needed relatively little effort to carry out and were easily maintained over time [17, 26] because changes in the IT-systems are permanent as well as independent of conditions in the individual clinics. However, as it is system-dependent, it may not be possible to apply everywhere.

The one study that used pop up alerts achieved a high reduction in number of tests requested [28]. In hospitals, this approach has been extensively used, both by applying non-interruptive alerts, which passively displays information, such as redundant testing or suggestions of corollary orders [28, 42], or interruptive alerts which require the user to acknowledge the alert [43, 44]. A theoretical benefit of these alerts is that they can provide “just in time” advice and thereby provide long-term education. Conversely, this type of interventions contain the risk of “alert fatigue” when the alert is displayed for long time periods, making the requester “blind” to the content of the alert and/or even irritated by the interruptions caused by the alert. This is probably most often seen, when alerts are used with tests that are requested in high numbers.

The use of reflexive tests or gate keeping rules have just recently been investigated among GPs (Table 3) both in regards to over-used and under-used tests and this intervention seem to achieve positive effects. The use of gate keeping whether it was combined or not with educational components obtained comparable results in the included studies, but we did not find a clear effect when reflexive testing was combined with education. Nevertheless, these advanced tools are described by the authors as interventions which are more suited when implementing clinical guidelines [19]. This is because laboratory professionals are in the position to manage demand on laboratory tests by utilizing evidence based guidelines, when developing specific test ordering directives and gate keeping rules [26].

According to the review of Cadamuro et al. [33], laboratory diagnostic algorithms are the most sophisticated tools to influence the appropriate use of diagnostic testing as they permit physicians to obtain relevant laboratory results without ordering individual tests and still arrive at a definitive diagnosis. It has been stated that this approach together with a comprehensive, patient-specific interpretation of the test results, can substantially reduce the cost by decreasing the time to diagnosis and the number of tests ordered without compromising patient care [39–41]. CPOE approaches are still only sparsely investigated in the primary care setting, however the studies included in this review are mainly medium risk of bias, suggesting that additional low risk of bias investigations are needed in the future.

In a number of studies, interventions were designed in collaboration and applied in consensus with the requestors [9, 10, 12, 16–18, 21, 27, 28, 31]. In some of this [16, 20, 28] the opinion of the GPs as to the applied interventions was investigated by questionnaires or interviews. However, none of the studies reported in this review have evaluated
how the GPs view the impact of the different interventions on their daily practice. GPs are exposed to the intervention every day and, in the perspective of improving the laboratory test requests, having their opinion is essential. If the strategy applied is considered annoying and/or delaying the workflow by the GPs, the applied strategy should be re-evaluated. The tendency of applying strategies in close collaboration with the requestors seems to be an effective method to share information as well as potential problems that otherwise may have been overlooked, but based on the present results, we cannot substantiate this. Nevertheless, we think that a mutual discussion may bring better results or at least better implementation, and further studies in this direction is recommended.

Included in the limitations reported, there were short follow-up periods and poor study designs. However, the most commonly mentioned limitation was regarding the effect of the interventions on the patients. In all intervention studies, it was stated that it was unknown whether the reduced number of tests after intervention might have led to a missed diagnosis in patients. Evaluating appropriateness of test requests both before and after applying an intervention is a difficult task, especially when the clinical records of the patients are not integrated in the laboratory system. But, if the intervention is well designed and the appropriate use of the test under investigation is well documented and subject to clear clinical guidelines, changes in number of tests requested ought not to affect patient care.

**Costs**

The studies which reported changes in costs either as marginal costs or costs of reagents did not evaluate the cost-benefit of the interventions applied. Reporting only the reduction in costs without including any additional costs to the laboratory that might be necessary in regards to setting up and maintaining the intervention or informing the requestors, does not give a clear picture of the total expense involved. What could be done was to summarize the effect of how many blood test taken would potentially be taken if the intervention was not applied and consequently evaluate whether it was worthwhile to apply that intervention.

**Limitations**

Our review have some limitations. First of all, our search was only applied in two databases so potentially important studies included in other databases could be missed. In addition, the search terms did not include ‘demand management’ and it cannot be ruled out that some studies may have been missed on that account. Furthermore, because of the heterogeneity of the calculation of results in the different studies, we were not able to make a traditional meta-analysis, so our study is mainly descriptive. However, by adjusting the results, we were able to make the included studies comparable as to the relative changes in test numbers.

**Conclusions and future research direction**

This review gives an overview of the strategies recently applied in primary care to optimize use of laboratory tests as this setting is less well investigated compared to the hospital setting. We report that interventions including educational components consistently changed the number of tests and these results were supported by the good quality of the studies. Feedback reports have mainly been applied in combination with educational interventions, while when used alone the effect has been found to be minimal. The use of administrative changes, both alone and in combination with education seem to produce a marked change in the number of test requests, however the quality of those studies was generally considered to be of medium or high risk of bias, making the results less reliable. The use of soft and hard blocking were recently implemented in primary care, but the number of studies (mainly of medium risk of bias) was too low to extract clear conclusions.

The laboratory costs reported by some of the studies are supplementary information to the changes in number of tests and do not add further information to the analysis.

Future research is needed to evaluate the effectiveness of both soft and hard blocking as well as other approaches and relevant combinations of interventions that remain uninvestigated in primary care. For example, the use of feedback reports in combination with CPOE interventions. Furthermore, the minimum retesting intervals which have been extensively used in the hospital setting and where clear guidelines have been published [45] awaits proper investigation in primary care.

Designing and applying interventions in consensus with the GPs seems to be a good practice for sharing information between laboratory and primary care setting. However, as all the studies which used this method were lacking follow-up periods, it was not possible to say...
whether the applied interventions have resulted in long-lasting changes.

Moreover, it can be expected that collaboration on the optimization process will influence the GPs’ opinion of the interventions applied and the evaluation of appropriateness which should be included in future studies together with changes in number of tests.

Research funding: None declared.

Author contributions: All authors have accepted responsibility for the entire content of this manuscript and approved its submission.

Competing interests: Authors state no conflict of interest.

References