

Defining over- and underuse of laboratory diagnostics: a consensus-oriented E-Delphi study

*MSc Health Sciences
Faculty Science and Technology (TNW)
University of Twente,
Enschede, The Netherlands, 2020*

*First supervisor University of Twente
Committee Chair
Kusters, G.C.M, prof, dr.*

*Second supervisor
Committee member UT
Kip, M.M.A, dr.*



Defining over- and underuse of laboratory diagnostics: a consensus-oriented E-Delphi study

Rosalie Constance Visser

2020, Augustes

Faculty of Science and Technology, University of Twente, Enschede, the Netherlands

Abstract

Background: Laboratory diagnostics are essential for the detection, management, and monitoring of diseases. Most of the clinical decisions are facilitated by the test results of laboratory diagnostics. However, there is a reasonable risk that laboratory diagnostics may be used inappropriately. However, defining this concept is difficult as there is a general lack of consensus on how to define over- and underuse. This complexity is well demonstrated by the great variety in definitions used to classify laboratory diagnostics as being over- or underused, which highlights the need for a clear definition.

Objective: To reach a consensus on a definition of over- and underuse of laboratory diagnostics when taking into account the perspectives of the user, provider and, payer in primary and secondary care.

Design: Literature review and E-Delphi.

Population: Thirteen participants from three different fields of expertise: clinicians (users) who routinely request laboratory diagnostics (n = 4), clinical chemists (providers) who analyze and provide laboratory test results (n = 4), and health insurers (payers) involved in the reimbursement of laboratory diagnostics (n = 5).

Methods: A literature review was conducted on published studies from January 2000 to March 2020. Studies on over- and underuse of laboratory diagnostics, ranging from blood and urine samples to tumor markers, were assessed. An E-Delphi was performed, which comprised four subsequent rounds of online surveys designed to capture the collective opinions of experts.

Results: Overuse was defined as ‘the use of diagnostic tests without any underlying reason or when the outcome is not relevant for the clinical question or decision-making, for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits’. Underuse was defined as ‘the failure to use diagnostic tests when this could have yielded relevant findings for the decision-making process, or for setting a correct and accurate diagnosis’.

Conclusion: The definitions made by the expert panel during this E-Delphi study can be used to facilitate the use of uniform terminology when investigating or discussing over- and underuse of diagnostic tests and enhance the use of common language across different fields of expertise. Still, it is essential to note that both definitions reflect a consensus within this E-Delphi study. They are not necessarily the ‘best’ or ‘correct’ findings, and considerations must be given to the validity and reliability which challenge the performance of this E-Delphi study.

Keywords: Definition, E-Delphi, systematic literature review, laboratory diagnostics, overuse, underuse, inappropriate use, appropriateness, primary- and secondary care.

Topic: Over- and underuse of laboratory diagnostics

1. Introduction

Laboratory testing is essential for the detection of diseases, and for the management and monitoring of previously diagnosed conditions (1, 2). Besides, most of the clinical decisions are facilitated by the test results of laboratory diagnostics. It is roughly

estimated that the results of laboratory diagnostics contribute to 70% of clinical decisions (1, 3). Thereby, laboratory medicine has become a vital discipline in modern healthcare (3).

As being a vital discipline contributing to high-value and high-quality medical outcomes, there is a reasonable risk that laboratory diagnostics may



be used inappropriately (1). This is partly due to (i) the high rate of technological innovation, (ii) its crucial role in many conditions and clinical settings, (iii) the increasing frequency in which laboratory diagnostics are being ordered, and (iv) the growing pressure to reduce costs and waste of resources, increase the quality of care, and improve the effectiveness of care and patient outcomes (3, 4). Following this, over- and underuse can lead to delayed, missed, or incorrect diagnoses and adversely affects the patient's treatment process and increases costs for healthcare (1, 3, 5). Hence, laboratory medicine is an area of special interest for the investigation of inappropriateness (4).

Inappropriate testing does not only include the excessive use of tests (overuse), but also the failure to order the appropriate tests (underuse) (3). Literature indicates that inappropriate testing, further referred to as over- and underuse, ranges from 11% to 70% for general biochemistry and hematology tests, 5% to 95% for urine screens and microbiology, and 17.4% to 55% for cardiac enzymes and thyroid tests (6). So, although variations in rates of inappropriate testing occur, categorizing this practice as being over- or underuse can be more complex. This may be because over- and underuse includes four basic principles, effectiveness, efficiency, ethical justifications, and relevance, each of which has a different meaning for the patient, clinician, payer, healthcare system, and society (7).

This complexity is also demonstrated by the great variety in definitions used to classify laboratory diagnostics as being over- or underused and their general lack of consensus (8, 9). Therefore, this current research aims to investigate expert opinions to reach a consensus on the definition of over- and underuse. Hence, the following research question is formulated: *'How can over- and underuse of laboratory diagnostics be defined, when taking the perspectives of the user, provider, and payer in primary and secondary care into account?'* This current research attempts to answer the question by conducting a literature review and by performing an E-Delphi study.

2. Methods

Literature review

Study design

The literature review was conducted according to the five-phases of literature review of Khan, *et al* (2003) consisting of (i) framing questions for analysis, (ii) identifying relevant articles, (iii)

assessing the quality of articles, (iv) summarizing the evidence, and (v) interpreting the findings (10). The purpose of this literature review was to explore existing literature on over- and underuse of laboratory diagnostics, to be able to provide for a theoretical framework for the E-Delphi surveys. Therefore, the following research questions were formulated:

1. How is over- and underuse conceptualized in literature?
2. What definitions is/are currently used for over- and underuse?
3. What are the consequences of over- and underuse towards patient safety and financial burden on society?

Data source and search strategy

PubMed and Scopus were used to search for articles describing over- and underuse of laboratory diagnostics published between 2000/01/01 and 2020/12/31. The literature search was conducted in the period of March 2020. The following search terms and synonyms were combined: "inappropriate use", "overuse", "underuse" "laboratory diagnostics", "laboratory utilization", and "consequences". Appendix I provides an overview of the comprehensive search strategy and of the combinations of search terms used.

Eligibility criteria

Articles that reported about over- and underuse of laboratory diagnostics were selected, regardless of the study design used. We considered all laboratory diagnostics ranging from blood and urine samples to tumor markers. Articles were included when they had been found relevant to one or more of the three questions mentioned above and when they were written in English or Dutch. Studies were excluded when they only focused on diagnostics other than laboratory diagnostics.

2.2 E-Delphi study

Study design

The E-Delphi study was conducted according to the recommendations described by Hasson, *et al* (2002) on how to use the Delphi survey technique (11). This E-Delphi study aims to investigate expert opinions to reach a consensus on the definition of over- and underuse of laboratory diagnostics.



Panel selection

We compiled an expert panel using a non-probability purposive sampling technique combined with the quota technique. In total, 12 participants were invited via a personal email. All participants were specifically approached based on their expertise. We purposely aimed to include participants from three different fields of expertise: (i) clinicians (users) who routinely request (> 3 times a week) laboratory diagnostics in primary and secondary care, (ii) clinical chemists (providers) who analyze and provide laboratory test results to assist clinical decision-making, and (iii) health insurers (payers) involved in the reimbursement of laboratory diagnostics. Patients (consumers) were not involved in this E-Delphi study since it is almost impossible for the individual patient to recognize or assess over- and underuse of laboratory diagnostics (8).

Development of the questionnaires

This E-Delphi study comprised four subsequent rounds of online surveys. The surveys were developed based on a review of the literature whereof the four basic principles of the appropriate use of laboratory diagnostics which include effectiveness, efficiency, ethical justification, and relevance were used as theoretical framework (7). In the surveys and results of the E-Delphi study, the term ‘diagnostic tests’ will be used to refer to laboratory diagnostics. The first survey consisted of seven open questions, of which four questions focused on the four basic principles, and three questions focused on over- and underuse in practice. The subsequent surveys consisted of closed-ended questions and were based on the responses given in the previous round. Each survey was piloted by at least two research professionals to check for clarity. Ethical approval for this study was granted by the Behavioral, Management, and Social sciences (BMS) Ethics Committee of the University of the Twente in the Netherlands on 17/04/2020 (No. 200433).

Data collection

The first survey was distributed around mid-April, the subsequent questionnaires at a three weeks interval until the beginning of July. Appendix II provides a comprehensive overview of the timetable of the E-Delphi study. Each questionnaire was designed and distributed using Qualtrics. Participants received a personal email with an anonymous reusable link to each survey, followed by two or three reminder emails, at a one-week interval, per round. Participants were

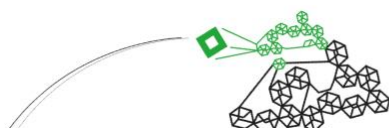
requested to provide online informed consent. Analysis of the qualitative data was conducted by thematic analysis in AtlasTI 8.4 and analysis of the quantitative data by descriptive statistics in Excel. In this E-Delphi study, a consensus was defined as a percentage agreement of at least 65% amongst all participants depending on the round of the survey. Details are given in the following sections.

Round 1 (divergence)

Participants were requested to provide their views on the four basic principles (effectiveness, efficiency, ethical justification, and relevance) and on over- and underuse in practice to capture the collective opinion of experts. The expert panel was allowed to provide as much information as they wished within the free text space provided. In the first round, participants were also requested to provide some socio-demographical characteristics including age, gender, employment status, and time (in years) employed in their field of expertise. Responses of the open-ended questions were thematically grouped into joint statements, merging similar topics which were described using different terminology.

Round 2 (convergence)

In this round, the survey consisted of the same seven questions as in round one, but instead of open-ended questions, each question entailed of a list of possible statements collected from the first round. Participants were requested to indicate whether they agree or disagree with a statement on a Likert scale of 1 to 5 (1 strongly disagree – 5 strongly agree). A consensus was reached if more than 65% of the participants rated the statement with a three or higher on the Likert Scale. The first round also allowed the expert panel to indicate whether they considered that the candidate definitions should relate to all four basic principles (effectiveness, efficiency, ethical justification, and relevance) or a selection. Therefore, we requested participants to indicate the importance of a specific basic principle being part of the candidate definitions on a Likert scale of 1 to 5 (1 being not important – 5 being very important). A consensus was reached if more than 90% of the participants rated the basic principle with a three or higher on the Likert Scale. So, after this round, the expert panel determined which basic principles were important enough to be included in the candidate definitions and which statements were used as input for the third survey.



Round 3 (convergence)

The third round focused specifically on the parts that comprised the preferred statements collected from the second round. Therefore, each statement was disaggregated into two, three, four, or more aspects. Participants were requested to indicate whether they agree or disagree with the aspects on the same Likert Scale of 1 to 5 (1 strongly disagree – 5 strongly agree). A consensus was reached if more than 90% of the participants rated the aspect with a three or higher on the Likert Scale. After this round, the expert panel has thus determined which essential parts of the statements they preferred to be included in the candidate definitions. These aspects were used as input for the fourth round.

Round 4 (convergence)

The basic principles, statements, and aspects together represented the unique attributes of the candidate definitions. The first candidate definition both for over- and underuse consisted of one aspect, to which several aspects, were subsequently added. Participants were requested to rate these candidate definitions on a Likert scale of 1 to 10 (1 extremely disagree – 10 excellent). The analysis of the total sum of ratings, median, and percentage agreement were combined to assess if one of the definitions would be preferred by the majority of the expert panel.

3. Results

Results of the literature review

Search results and characteristics

The search strategy resulted in 372 hits on Scopus, and 66 on PubMed, thereof 19 duplicates were removed, and ultimately resulted in 419 unique articles. Based on the title and abstract, 29 studies were initially eligible. After reading the full-text and applying the inclusion criteria, 16 publications were included. The assessment of the references lists resulted in the inclusion of 4 additional articles. So, in its entirety, 20 studies were included (Fig. 1). The characteristics of the included studies are shown in Table I.

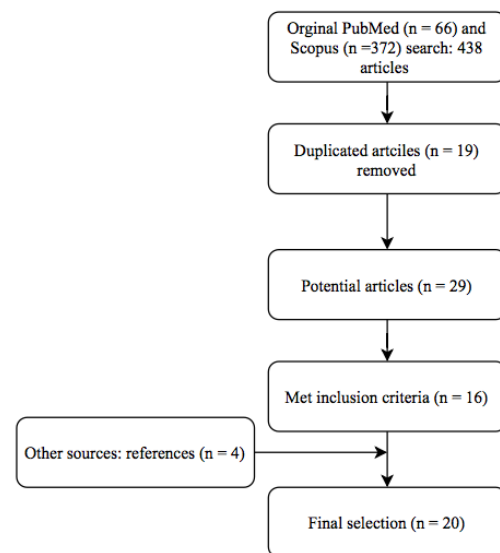
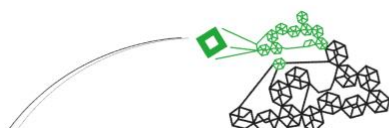


Fig. 1 Search results

Table I Characteristics of included articles

	N (%)
Study design	
Review article	7 (35%)
(Retrospective) case study	6 (30%)
Meta-analysis	2 (10%)
Systematic literature review	2 (10%)
Realist review	1 (5%)
Research article	1 (5%)
Opinion paper	1 (5%)
Domain	
Over- and underuse	12 (60%)
Overuse	7 (35%)
Not specified	1 (5%)
Underuse	0 (0%)
Relevant to sub-question 1	
Yes	20 (100%)
No	0 (0%)
Relevant to sub-question 2	



Yes	13 (65%)
No	7 (35%)
Relevant to sub-question 3	
Yes	15 (75%)
No	5 (25%)

Definitions of over- and underuse and (in)appropriateness

Although over- and underuse of diagnostic tests commonly occurs in practice, defining this concept is difficult, and gets even more difficult when over- and underuse occur simultaneously (8). The complexity of over- and underuse is well demonstrated by its wide variety in definitions (7, 9). Most authors define over- and underuse by estimating the appropriateness through various audit criteria like adherence to guidelines endorsed by an organization, primary literature, local consensus, rules developed and validated within research articles, or individual opinions (9, 12-16).

In 1986, the worldwide RAND (research and development) Corporation defined appropriateness as ‘the expected health benefit exceeds the expected negative consequences by a sufficiently wide margin that the procedure is worth doing, excluding considerations of monetary costs’ – focusing on the safety of the individual patient (7). In 1996, avoiding unnecessary waste of resources and potential medical errors increases in importance and according to the College of American Pathologists appropriateness relates to ‘the extent to which a particular procedure, treatment, test or service is effective, clearly indicated, not excessive, adequate in quantity, and provided in the inpatient, outpatient, home or another setting best suited to the patient’s needs’ (2, 7). Two years later, van Walraven and Naylor suggest that inappropriate testing involves any test that could reasonably be avoided at no significant detriment to a patient’s care’ (7). Following this, definitions of appropriateness written by (international) healthcare organizations focus on patients’ preferences and safety – ‘the degree to which service is consistent with a clients’ expressed requirements and is provided in accordance with current best practice’ or ‘the degree to which the care and services provided are relevant to an individual’s clinical needs, given the current state of knowledge’ (7). A few years later, around 2003, a development related to shifting from outcome-based to evidence-based medicine allowed for a new view regarding appropriateness focusing on clinical decision-making – ‘an

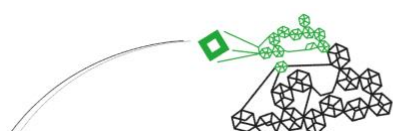
appropriate test is one in which the result provides an answer to a question that enables a decision to be made and an action taken’ (2, 7).

More recently, appropriateness is defined as ‘the outcome of a process of decision-making that maximizes net individual health gains within society’s available resources’ (4). Lippi (2015) mentioned a simple six R paradigm where an appropriate test can identify or rule out the presence of a given disease where the right test, with the right method, at the right time, to the right patient, at the right cost, for the right outcome is performed (17). Some other modern variations of inappropriate testing are ‘the use of tests which failed to be relevant to the signs and symptoms, diagnosis, monitoring of treatment, medical intervention, and confirmation of suspected conditions of the patient’ and ‘the use of tests that were unrelated to clinical indication and occurred when the blood sample was hemolyzed, results were not plausible, and/or the sample was inefficient’ (6). Furthermore, some authors defined over- and underuse separately as ‘a test is ordered but not necessary (overuse) and a test is necessary but not ordered (underuse) (12) or as a circumstance where a test not directly contributing to patient care is ordered (overuse) and as a circumstance where one or more tests needed for patient care are not ordered (underuse) (18).

Also, many authors defined over- and underuse by estimating the appropriate use of diagnostics as advised in guidelines (9, 12-16) – ‘practicing medicine in opposition to an organization guideline’ or ‘tests ordered in violations of a guideline produced by a government or professional society’ (9, 16). O’sullivan, *et al* (2018), defined over- and underuse as ‘a diagnostic test was ordered when the relevant guideline recommends not ordering it or a diagnostic test was not ordered when the relevant guideline recommends ordering it’ (14).

Quantifying over- and underuse

Besides the great variety in definitions, several studies have also observed a great variation in the quantification of over- and underuse in the primary- and secondary care (7-9, 12, 14, 15). A



15-year meta-analysis of Zhi, *et al* (2013), investigating 50 of the most common ordered diagnostic tests in the primary- and secondary care showed an overall mean rate (proportion) of overuse of 20.6% (95% CI 16.2% – 24.9%) and an overall mean rate (proportion) of underuse of 44.8% (95% CI 33.8% - 55.8 (12). Also, a meta-analysis of O’Sullivan investigating 47 diagnostic tests in primary care, showed that 17 tests were underused (17/47, 36%), and 11 tests were overused more than 50% of the time (11/47, 23%) (14). Miyakis, *et al* (2006) retrospectively evaluated 426 patient records in secondary care, whereby 25 diagnostic tests were assessed on clinical usefulness. As many as 68% (16,648/24,482) of these tests were found irrelevant for the clinical decision-making (19). More recently, a case study of Sarkar (2017) evaluated 200 patient records, also in secondary care, for the assessment of over- and underuse of diagnostic tests. Over three-quarters, of all cases (77,5%, 155/200) were associated with inappropriate testing, whereof 16% (32/200) with overuse, 44% (88/200) with underuse, and 17,5% (35/200) with both over- and underuse (20).

Patients’ safety issue

The consequences of over- and underuse towards patients’ safety are many and multifaceted (1, 4, 8, 16). The literature showed that over- and underuse both can result in delayed, missed, or incorrect diagnoses and to extended hospital stay (3, 4, 8, 12, 14, 21, 22). The consequences of underuse may be crucial independently of the rate of inappropriateness, as missing or delaying a diagnosis may be vital also if it occurs in a few or even in one patient (4). The consequences of overuse may result in unnecessary blood draws, anxiety, and stress for the patient; it also increases the possibility of false-positive outcomes which may lead to more unwarranted additional interventions being more invasive and harmful to the patient (6, 12, 19, 21).

Financial burden

Diagnostic tests contribute between a range from 1.2% to 5% of the public healthcare expenses and the amount less than 5% of the total hospital expenditures (4, 5, 12, 18). Even though diagnostic tests only contribute to a small

proportion of total healthcare expenditures, they exert a great influence on clinical decision-making (4, 5). A review article of Bogavac-Stanojevic (2017) on the cost-effectiveness of laboratory testing suggests that costs due to over- and underuse accounts for expenses higher than those due to the direct costs of appropriate testing (5). These additional expenses may arise from downstream activities such as unnecessary follow-up, waste of resources, and a prolonged hospital stay due to delayed, incorrect, or missed diagnoses (1, 3, 5, 12, 18). Furthermore, it is roughly estimated that 15% of hospital expenditures in the Organization for Economic Cooperation and Development (OECD) countries consist of treating quality of care failures (3). A case study of Sarkar (2017) has estimated the costs of a hospital with around 450 beds as a result of over- and underuse. Costs due to overuse were estimated at around 20.000 dollars and costs due to underuse were estimated equal to or more than 200.000 dollars, annually in the United States of America (20).

Results of the E-Delphi study

Participants

In total, 13 participants from three different fields of expertise were included in this E-Delphi study, 12 via purposive selection and one via snowballing. The expert panel consisted of five participants representing the payers, four that represent the providers, and another four that represent the users. The first survey was completed by 12 participants (92,3% response rate), 12 of 13 completed round 2 (92,3% response rate), 13 of 13 completed round 3 (100% response rate), and 11 of 13 completed round 4 (84,6% response rate). The demographic characteristics were collected for the 12 participants who participated in the first round. The majority of participants had been working in the field of expertise for more than five years, fulltime. For the clinicians and clinical chemists, the years employed in the field of expertise also includes the educational period. Gender distribution was evenly distributed. Table II provides a summary of the socio-demographic characteristics of the participants.

Table II The socio-demographic characteristics of the participants

Gender	N (%)
Male	58,3% (n = 7)
Female	41,7% (n = 5)
Age in years	



21 – 29 years	8,3% (n = 1)
30 – 39 years	16,7% (n = 2)
40 – 49 years	58,3% (n = 7)
50 – 59 years	16,7% (n = 2)
Years employed in the field	
5 – 9 years	33,3% (n = 4)
10 – 14 years	8,3% (n = 1)
15 – 19 years	33,3% (n = 4)
20 – 24 years	8,4% (n = 1)
25 – 29 years	16,7% (n = 2)
Employment status	
Fulltime	83,3% (n = 10)
Part-time	16,7% (n = 2)

Results round 1

The first round was completed by 12 participants, which provided their views on the four basic principles (effectiveness, efficiency, relevance, and ethical justification) and on over- and underuse in practice. Similar responses were grouped and compounded into joint statements through thematic analysis. The thematic analysis resulted in the development of thirty-seven joint statements, that is, twenty-one statements representing the basic principles and sixteen statements representing over- and underuse in practice. Most responses that were included in the second survey remained unchanged except for the addition of a few verbs and prepositions to improve readability. Appendix III provides the design of the first survey.

Results round 2

Participants were presented with the four basic principles and thirty-seven statements on the four

basic principles and on over- and underuse in practice. The expert panel was requested to indicate whether they agreed or disagreed with the statements and to indicate the importance of a specific basic principle being part of the candidate definitions. Participants' responses showed an agreement for 22 of the 27 statements of at least 65% (three representing efficiency, three representing relevance, four representing overuse, five representing underuse, and four representing correct use). The basic principles, effectiveness, efficiency, and relevance were considered sufficiently important to be included in the candidate definitions (>90%), ethical justification, was not further included as it only reached an agreement of 67% (<90%) (see table III). Table IV and V show the statements with the highest percentage agreements. Appendix IV provides the design and results of the second survey.

Table III Qualitative results of round 2 – the four basic principles

	Median	Range	Percentage agreement
Effectiveness	5	4 – 5	100%
Efficiency	4	3 – 5	92%
Relevance	4	3 – 5	92%
Ethical justification	4	3 – 5	67%

Table IV Results of round 2 – basic principles (effectiveness, efficiency, relevance)

	Median	Range	Percentage agreement
Effectiveness			
The use of diagnostic tests that are necessary for setting a diagnosis as correctly and quickly as possible, avoiding unnecessary and duplicate diagnostic tests whenever possible.	4	1 – 5	67%
Performing medically necessary diagnostic tests on a patient, to set the most likely and accurate diagnosis.	4,5	3 – 5	83%
The right test at the right time for the patient, in which the requester will adjust his treatment plan based on the results, and	4	2 – 5	75%



that other healthcare providers (at referral) can continue on this basis.

Efficiency

Performing a medically necessary diagnostic test on a patient, at the right time with the purpose of setting a diagnosis as quickly as possible, using as few resources as possible and minimizing the burden to the patient. 4,5 3 – 5 92%

Diagnostic tests that are used sensibly and effectively, in which not only the efficiency of the diagnostic method is relevant, but also the use of diagnostic tests provides an efficient and meaningful answer to the concerning issue. 4 1 – 5 67%

The use of diagnostic tests is effective, targeted, cost-efficient, and thus optimal use of effort and resources concerning the expected diagnostic gain. 4 3 – 5 67%

Relevance

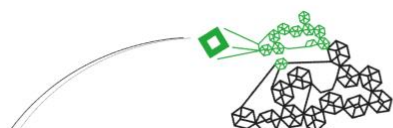
The problem-driven requesting of correct and relevant diagnostic tests resulting in findings that are important for setting a diagnostic conclusion and for the concerning question or decision-making. 4 2 – 5 83%

Diagnostic tests that can differentiate between different diagnoses or support treatment choices that are useful for the diagnosis, health monitoring, treatment, and/or prognosis of the patient. 4 3 – 5 83%

Performing the correct diagnostic tests on a patient, to set a diagnosis, provided that undergoing the tests outweighs setting a diagnosis and subsequently initiating meaningful, targeted, and effective medical treatment. 4 2 – 5 83%

Table V Results of round 2 – correct-, over- and underuse in practice

	Median	Range	Percentage agreement
Correct use			
If diagnostic tests are effective and ethical, and taking into account a good use of resources and whereby relevant information to set a correct diagnosis is obtained.	4	3 – 5	92%
If exactly those diagnostic tests have been performed that support the correct diagnosis, the monitoring of the patient's health, or the safe and targeted performance of treatment.	4	3 – 5	75%
When there is effective and efficient requesting behavior, where the correct diagnostic test is requested at the right time for the patient.	4	2 – 5	67%
When there is effective, efficient, ethical, and relevant use of diagnostic testing where relevant information is obtained to make policy and where there is no under- or overuse.	5	3 – 5	83%
Overuse			
When the results of a diagnostic test no longer contribute to the diagnosis or treatment policy.	4	3 – 5	83%
When diagnostic tests are performed without any underlying reason or based on improper grounds that do not provide an answer to the patient's health situation and is not necessary for the medical diagnostic process.	5	4 – 5	100%
When diagnostic tests are used when the burden and impact of undergoing diagnostic testing for an individual patient and/or the costs of diagnostic testing for society do not outweigh the added value of setting a diagnosis and the number of patients that can be treated more effectively after setting a diagnosis is limited.	4	2 – 5	67%



When diagnostic tests are not effective and/or socially justifiable, because too many resources are used and/or information is obtained that is not entirely relevant for setting a diagnosis or guiding a treatment.	4,5	3 – 5	83%
Underuse			
Failure to use diagnostic tests, while the burden and impact of undergoing diagnostic testing for an individual patient and/or the costs of diagnostic testing for society certainly outweigh the burden of setting a diagnosis and the number of patients that can be effectively treated after setting a diagnosis is large.	5	4 – 5	100%
When patient symptoms persist and diagnostic tests that are considered essential or highly desirable in (inter)national guidelines because these diagnostic tests can confirm or disprove a possible and probable underlying explanation, are not used.	4	3 – 5	75%
When diagnostic tests are omitted that could have yielded relevant findings and would have been effective and socially justifiable.	4,5	2 – 5	83%
Failure to use diagnostic resources or to request the right diagnostic test at the right time, while the result of such a resource is or may be relevant to the patient or his practitioner for diagnosis, treatment or prognosis.	4	4 – 5	100%
If too few diagnostic tests have been performed and too much uncertainty remains when setting a diagnosis or guiding a treatment, while a test is available that can reduce that uncertainty.	4	2 – 5	67%

Results round 3

The twenty-two preferred statements were disaggregated into 69 individual aspects, that is, 10 representing effectiveness, 12 representing efficacy, 13 representing relevance, 10 representing correct use, 14 representing overuse, and 10 representing underuse. Participants were presented with the aspects, and they were asked to indicate whether they agreed or disagreed with the aspects. Participants responses showed an

agreement for 14 of the 69 statements of at least 90% (three representing effectiveness, three representing efficiency, four representing relevance, one representing correct use, one representing overuse, and two representing underuse). Tables VI and VII show the results with the highest percentage agreements. Appendix V provides the design and results of the third survey.

Table VI Results of round 3 – basic principles (effectiveness, efficiency, relevance)

	Median	Range	Percentage agreement
Effectiveness			
Setting a correct diagnosis.	5	4 – 5	100%
Avoiding unnecessary diagnostic tests.	4	3 – 5	92%
Setting the most accurate diagnosis.	5	2 – 5	91%
Efficiency			
Diagnostic tests are used effectively.	4	3 – 5	92%
Diagnostic tests are used cost-efficiently.	4	3 – 5	92%
Optimal use of resources concerning the expected diagnostic gain.	4	4 – 5	100%
Relevance			
An outcome that is important for setting a diagnosis.	5	4 – 5	100%
An outcome that is important for the concerning question.	4	3 – 5	92%
An outcome that is important for the decision-making process.	5	4 – 5	100%
Diagnostic tests useful for patient prognosis.	4	3 – 5	91%

Table VII Results of round 3 – correct-, over- and underuse in practice

	Median	Range	Percentage agreement
--	--------	-------	----------------------



Correct use			
The right test at the right time.	4	3 – 5	92%
Overuse			
Performing diagnostic tests without any underlying reason.	5	4 – 5	100%
Underuse			
When diagnostic tests are considered essential or highly desirable in (inter)national guidelines, are not used.	4	3 – 5	92%
When diagnostic tests could have yielded relevant findings.	4	2 – 5	92%

Results round 4

The fourteen selected aspects were merged into 28 different candidate definitions, that is, 14 candidate definitions representing overuse, and 14 candidate definitions representing underuse. More specifically, the first candidate definition that was presented to the expert panel was formulated based on one aspect, and subsequently, the aspects were added one by one. The expert panel was requested to rate the candidate definitions with a maximum score of 10 per

definition, per participant. In total, 11 participants have rated the candidate definitions.

The final definition for overuse, see box I, was chosen as the best definition by 55% of the expert panel with a total sum of ratings of 85 (10 – 110) and with a median score of 8 (range 1 – 10). The final definition for underuse, see box II, was chosen as the best definition by 64% of the expert panel with a total sum of ratings of 86 (10 – 110) and with a median score of 7.5 (range 1 – 9). Appendix VI provides the design and results of the fourth survey.

Box I: Final definition for overuse.

Overuse is the use of diagnostic tests without any underlying reason or when the outcome is not relevant for the clinical question or decision-making, for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.

Box II: Final definition for underuse.

Underuse is the failure to use diagnostic tests when this could have yielded relevant findings for the decision-making process, or for setting a correct and accurate diagnosis.

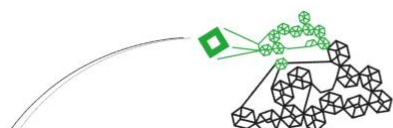
4. Discussion

4.1 Interpretation of findings

This research aimed to investigate expert opinions to reach a consensus on the definition of over- and underuse of laboratory diagnostics, and therefore, an E-Delphi study was performed.

A consensus was achieved on a definition for overuse as: ‘the use of diagnostic tests without any underlying reason or when the outcome is not relevant for the clinical question or decision-making, for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits’. Furthermore, a consensus was reached on a definition of underuse as: ‘the failure to use diagnostic tests when this could have yielded relevant findings for the decision-making process, or for setting a correct and accurate diagnosis’.

Within this E-Delphi study, a consensus was defined as a percentage agreement of at least 65% amongst all participants depending on the round of the survey. In practice, the most common definition for a consensus is the percentage agreement which is generally based on sample size, the aim of the study, the available resources, or it is decided on ad-hoc, as the literature provides few guidelines (11, 23, 24). Hence, it is essential to note that both definitions reflect a consensus within this E-Delphi study and that they are not necessarily the ‘best’ or ‘correct’ findings. Furthermore, during the analysis, we observed an increase in percentage agreement across rounds two and three, but a decline in percentage agreement in the last round. Across round two and three, participant strongly agreed on some specific aspects which should be included in the candidate definitions. Following this, these stand-alone aspects were merged into



candidate definitions and presented in the last round. In consequence, participants may agree more with one part of the definition than with the other, which resulted in a decrease of percentage agreement in the last round. Thereby, the percentage agreement of the candidate definitions also showed little differences in the final round. Therefore, a consensus of the definition on over- and underuse was also dependent on the median, range, and total sum of ratings.

4.3 Strengths and limitations

This E-Delphi study has several strengths. Foremost, the strength includes that this research is based on the hypothesis that mutual knowledge enriches individual judgments and captures the collective opinions of experts. In the beginning, participants were allowed to provide as much information as desired to maximize the possibility of exposing the most vital judgments. No guidance from literature or previous studies of possible answers were offered in order not to influence the participant's opinions. The analysis was performed with high cautiously to avoid unintentional modification of participants' responses and deviation from the originally intended meaning. Thus, responses remained unchanged except for the addition of a few verbs and prepositions to improve readability. Participants identities were anonymized, and the online surveys were independent so that each opinion carries equal importance and weight during the analysis. Furthermore, participants may feel more comfortable giving judgments on certain topics in anonymous form. Last, participants have shown a high commitment within this E-Delphi study as the drop-out rate was low across all four rounds.

Although the E-Delphi method is widely accepted in healthcare research, it is also criticized, as issues regarding validity and reliability challenge the performance of an E-Delphi study (11, 25). Therefore, there are, some limitations that need to be recognized. The expert panel was not selected randomly but via purposive sampling. Hence, representativeness may not be assured as this have led to including only participants who are interested in the topic on over- and underuse of diagnostic tests. Accompanied this by the sample size, which may be too small to be representative of all three fields of expertise (11). Following this, considerations must be given to the reliability, as if the same surveys were given to other expert panels, there may be a low chance that similar findings will be obtained. Also, across rounds

participants may have changed their views, and move towards consensus because other participants have identified a more relevant topic that they had not considered. Therefore, validity may be compromised due to participants changing individual judgments as a result of a different view from the majority of the expert panel (23). Besides, within the E-Delphi study, the expert panel had no opportunity to discuss or elaborate on the issues raised. The missing face-to-face meetings impede the participants from deliberating essential topics, such as clarifying terms used and argue differences in opinions. Hence, terms as 'correct diagnosis', 'accurate diagnosis', or 'concerning question' were not further specified and, therefore, free to interpret.

4.4 Comparison with existing literature

Although literature acknowledges that over- and underuse of diagnostic tests commonly occurs in practice, defining this concept is difficult since it includes different basic principles (effectiveness, efficiency, ethical justifications, and relevance), different meanings for the patient, clinician, payer, healthcare system, and society, and the possibility of simultaneous occurrence (7, 8). There have been several attempts to define over- and underuse and previous studies usually rely on determining the appropriate use of diagnostic tests as advised in guidelines endorsed by an organization, primary literature, local consensus, rules developed and validated within research articles, or individual opinion's (9, 12-16). Published guidelines have become increasingly popular as audit criteria, and despite some well-known limitations, including varying quality of guidelines due to a lack of evidence or consensus, the accepted reference of defining over- and underuse is through adherence to guidelines (7, 9, 14). The considerable variability in definitions and the various audit criteria used to classify diagnostic tests as being over- or underused, and their general lack of consensus highlighted the need for a clear definition (7, 9). During this E-Delphi study, we reached a consensus on the definition of over- and underuse based on the basic principles of the appropriate use of laboratory diagnostics (effectiveness, efficiency, ethical justification, and relevance), while taking into account the perspectives of the user, provider, and payer.

4.5 Implications for research and practice

The added value of this E-Delphi study is that the definitions might facilitate the use of uniform terminology when investigating or discussing



over- and underuse of diagnostic tests and enhance the use of common language across different fields of expertise. Research designed to refine the definitions, identified by the expert panel, in different contexts, and across different stakeholders would be helpful. Furthermore, we acknowledge that for care to be patient-centered, patients need to participate in the research that informs healthcare decisions. Therefore, future studies should focus on an even broader multidisciplinary methodology and also take into account the perspective of the patient.

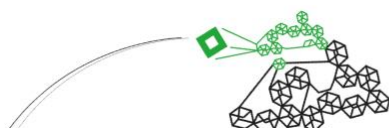
5. Conclusion

This E-Delphi study emphasized the need for a clear definition of over- and underuse of laboratory diagnostics. Given that there is considerable variability in definitions and audit criteria used to classify diagnostic tests as being over- or underused, which is accompanied by a general lack of consensus. The high commitment of the participants enabled the expert panel to

reach a consensus on the definition of over- and underuse of laboratory diagnostics within four rounds of online surveys. The definitions made by the expert panel during this E-Delphi study can be used to facilitate the use of uniform terminology when investigating or discussing over- and underuse of diagnostic tests and enhance the use of common language across different fields of expertise. Still, it is essential to note that both definitions reflect a consensus within this E-Delphi study. They are not necessarily the 'best' or 'correct' findings, and considerations must be given to the validity and reliability which challenge the performance of this E-Delphi study.

Acknowledgment

The author thanks all experts for their participation. Special thanks to the first supervisor prof. dr. G.C.M. Kusters, and second supervisor dr. M.M.A. Kip from the University of Twente for their valuable contributions and feedback.



Appendix I Search terms and strategy

Several terms and synonyms were combined to find relevant articles describing the inappropriate use of laboratory diagnostics. PubMed and Scopus were used to search for articles from 2000/01/01 through 2020/12/31 in the period from the beginning of February to the end of March 2020. Additional filters were publication year > 1999, language limited to English and humans.

Query string Scopus

(ALL("inappropriate use" OR overuse OR overutilization OR underuse OR underutilization OR appropriateness OR inappropriateness) AND ALL("laboratory diagnostic" OR "laboratory medicine") AND ALL("laboratory utilization" OR "laboratory testing") AND ALL(consequences OR effects OR results OR concerns OR "patient safety" OR "financial burden")) AND PUBYEAR > 1999 AND (LIMIT-TO (LANGUAGE,"English")) AND (LIMIT-TO (DOCTYPE, "ar") OR LIMIT-TO (DOCTYPE, "re"))

Query string PubMed

((("Inappropriate use" OR overuse OR overutilization OR underuse OR underutilization OR appropriateness OR inappropriateness) AND ("laboratory diagnostic" OR "laboratory medicine")) AND ("Laboratory utilization" OR "laboratory testing")) AND (consequences OR effects OR results OR concerns OR "patient safety" OR "financial burden")



Appendix II Timetable of the E-Delphi process

<i>Date</i>	<i>Event</i>
25 th March – 15 th April 2020	Inviting participants to participate in the E-Delphi study (n = 12) + snowballing (n = 1)
20 th April 2020	Sent information letter + distribution of the first survey
27 th April 2020	First reminder sent out
4 th May 2020	Second reminder sent out
9 th – 16 th May 2020	Analysis of round 1 results and design of the new survey (dataset n = 12)
18 th May 2020	Distribution of the second survey (n = 13)
25 th May 2020	First reminder sent out
28 th May – 6 th June	Analysis of round 2 results and design of the new survey (dataset n = 12)
8 th June	Distribution of the third survey (n = 13)
17 th June	First reminder sent out
23 rd June	Second reminder sent out
25 th June	Third reminder sent out
27 th June – 2 nd July	Analysis of round 3 results and design of the new survey (dataset n = 13)
2 nd July	Distribution of round 4 survey (n = 13)
7 th July	First reminder sent out
13 th July	Second reminder sent out
15 th July	Analysis of round 4 results (dataset n = 11)



Appendix III Delphi questionnaire round 1 – 20th April 2020

Dear participant,

Following your participation in the study 'defining of over- and underuse of laboratory diagnostics' you are given the first survey. This study aims to investigate expert opinions to reach a consensus on the definition of over- and underuse.

In this first survey, you are presented with seven open-ended questions. The first four questions are based on the four basic principles of the appropriate use of laboratory diagnostics (effectiveness, efficiency, ethical justification, and relevance). Literature mentions these four basic principles as the starting point for the correct use of laboratory diagnostics. As an expert, you will be asked for your view on the four basic principles. In the following three questions, you will be asked for your opinion on over- and underuse in practice. By your opinion and that of the expert panel, possible statements of a candidate definition will be compiled.

As it concerns your opinion, there are no correct or incorrect answers. Answering this first survey will take about 15 minutes. Before you start the first survey, we ask you to confirm your participation in an online consent form. At the end of the survey, you will be asked to fill in five socio-demographic characteristics.

Should you have any further questions regarding the survey, your participation, or this questionnaire, please contact us via e-mail: xxx or mob: xxx

Q1 I have read and understood the information letter and consent form and consent to everything contained herein.

- I give consent
- I do not give consent

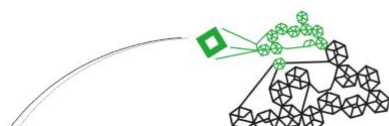
Q2 What do you understand by the effective use of diagnostic tests?

Q3 What do you understand by the efficient use of diagnostic tests?

Q4 What do you understand by ethical justifiable use of diagnostic tests?

Q5 What do you understand by the relevant use of diagnostic tests?

Q6 When do you think there's been overuse of diagnostic tests?



Q7 When do you think there's been underuse of diagnostic tests?

Q8 When do you think there's been correct use of diagnostic tests?

Q9 What is your gender?

- Female
- Male
- Otherwise

Q10 What is your age?

- < 21 years
- 21 - 29 years
- 30 – 39 years
- 40 – 49 years
- 50 – 59 years
- > 59 years

Q11 What is your current job position?

Q12 How long do you practice your profession? (including education period)

- 5 – 9 years
- 10 – 14 years
- 15 – 19 years
- 20 – 24 years
- 25 – 29 years
- > 30 years

Q13 Which of the following categories best describes your employment situation?

- Fulltime
- Part-time
- Retired
- Otherwise



Appendix IV E-Delphi survey round 2 – 18th May 2020

Dear participant,

Following your participation in the study 'defining of over- and underuse of laboratory diagnostics' you are given the second survey. This study aims to investigate expert opinions to reach a consensus on the definition of over- and underuse.

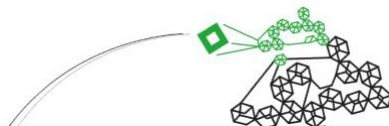
In the first round, 12 of 13 participants completed the survey. All answers were carefully discussed and analyzed by the research team. The responses were reduced to 37 statements using thematic analysis.

In this second survey, you are requested to indicate how important you consider a specific component or statement being part of the final definition on a Likert scale of 1 to 5 (1 = strongly disagree - 5 = strongly agree). Any component that is considered moderately important (>3) by at least 90% of the participants and any statements that are considered moderately important (>3) by at least 65% of the participants are added to the third survey.

As it concerns your opinion, there are no correct or incorrect answers. Should you have any further questions regarding the survey, your participation, or this questionnaire, please contact us via e-mail: xxx or mob: xxx

Q1 What do you understand by the effective use of diagnostic tests?

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
The use of diagnostic tests that are necessary for setting a diagnosis as correctly and quickly as possible, avoiding unnecessary and duplicate diagnostic tests whenever possible.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Performing medically necessary diagnostic tests on a patient, to set the most likely and accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When the diagnostic resources used, and the results that follow therefrom, are used to guide the diagnosis, treatment, or prognosis of a patient or patient group.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The right test at the right time for the patient, in which the requester will adjust his treatment plan based on the results, and that other healthcare providers (at referral) can continue on this basis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The effective use of the concerning diagnostic tests for the correct clinical purposes and clinical question, to achieve the intended effect.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



The correct diagnostic tests are requested, according to the guidelines, for the intended purpose of specifically answering a diagnostic question.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests of which each step or test is relevant to set a diagnosis or to allow treatment to be targeted and safe.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q2 Component: effectivity

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
How important do you think that 'effectivity' as a component is part of the final definition of over- and underuse?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q3 What do you understand by the efficient use of diagnostic tests?

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
Performing a medically necessary diagnostic test on a patient, at the right time to set a diagnosis as quickly as possible, using as few resources as possible and minimizing the burden to the patient.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of a limited number of requests and determinations to answer the referrer's clinical question, to set a diagnosis, and to determine any follow-up policy, where the aim must be that aim to improve the patient's health.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that are used sensibly and effectively, in which not only the efficiency of the diagnostic method is relevant, but also the use of diagnostic tests provides an efficient and meaningful answer to the concerning issue.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of diagnostic tests is effective, targeted, cost-efficient, and thus optimal use of effort and resources concerning the expected diagnostic gain.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q4 Component: efficiency

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
How important do you think that 'efficiency' as a component is part of the final definition of over- and underuse?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Q5 What do you understand by the relevance of diagnostic tests?

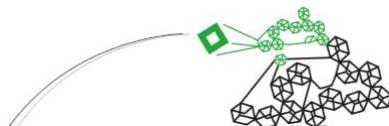
	Strongly agree	Disagree	Neutral	Agree	Strongly agree
The problem-driven requesting of correct and relevant diagnostic tests resulting in findings that are important for setting a diagnostic conclusion and for the concerning question or decision-making.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that can differentiate between different diagnoses or support treatment choices that are useful for the diagnosis, health monitoring, treatment, and/or prognosis of the patient.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When the correct diagnostic tests are used for the correct (composition) of complaints, not too many and not too few, and with a view to and input from the interests of the patient on which policy can be made that contributes to the diagnosis, treatment, and prognosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Performing the correct diagnostic tests on a patient, to set a diagnosis, provided that undergoing the tests outweighs setting a diagnosis and subsequently initiating meaningful, targeted, and effective medical treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q6 Component: relevance

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
How important do you think that ‘relevance’ as a component is part of the final definition of over- and underuse?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q7 What do you understand by the ethical use of diagnostic tests?

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
Performing sensible diagnostic testing of a patient that does not cause harm to the patient and which has added value for the individual, taking into account his or her life expectancy and life wishes, and taking into account the social health care budget.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Diagnostic tests that are performed in an equitable and scientifically justified manner and only used necessary for setting a diagnosis that can contribute to answering the clinical question.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The added value of diagnostic testing to the possible diagnosis or treatment justifies the possible outcomes of testing and the adverse impact that diagnostic tests may have on the patient or society.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
That the diagnostic tests used should provide an advantage for the patient and only detect something that leads to serious consequences in a reasonable period, or something that can be addressed to prevent later complications and unfavorable outcomes.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The responsible use of certain resources for a diagnostic purpose, whereby anyone who has the right or necessity for a test can obtain it and that this is not compromised by wasting these tests on irrelevant diagnostics.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The right test, with the right intention at the right time for the patient based on the expertise of the requester with transparency about how the diagnostic tests are performed and about the costs, when diagnostic tests are used for other purposes, the patient should always agree.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q8 Component: ethical justification

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
How important do you think that 'ethical justification' as a component is part of the final definition of over- and underuse?	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q9 When do you think correct use of diagnostic tests occurs?

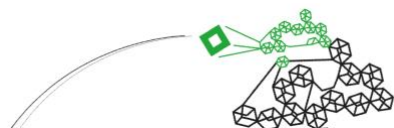
	Strongly agree	Disagree	Neutral	Agree	Strongly agree
If diagnostic tests are effective and ethical, and taking into account a good use of resources and whereby relevant information to set a correct diagnosis is obtained.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



If exactly those diagnostic tests have been performed that support the correct diagnosis, the monitoring of the patient's health, or the safe and targeted performance of treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When the current guidelines or standards have been correctly interpreted and followed, complemented by diagnostic reviewing, and the patient has been involved in the assessment and has been able to make an appropriate contribution to the decision-making process.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When there is effective and efficient requesting behavior, where the correct diagnostic test is requested at the right time for the patient.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When there is effective, efficient, ethical, and relevant use of diagnostic testing where relevant information is obtained to make policy and where there is no under- or overuse.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q10 When do you think overuse of diagnostic tests occurs?

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
When the results of a diagnostic test no longer contribute to the diagnosis or treatment policy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests are performed without any underlying reason or based on improper grounds that do not provide an answer to the patient's health situation and is not necessary for the medical diagnostic process.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests are used when the burden and impact of undergoing diagnostic testing for an individual patient and/or the costs of diagnostic testing for society do not outweigh the added value of setting a diagnosis and the number of patients that can be treated more effectively after setting a diagnosis is limited.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that are used without the patient's (or practitioner's) questioning or decision-making properly processed and without doing justice to the dilemmas that lie beneath each diagnostic assessment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests are not effective and/or socially justifiable, because too many resources are used and/or	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



information is obtained that is not entirely relevant for setting a diagnosis or guiding a treatment.

Q11 When do you think underuse of diagnostic tests occurs?

	Strongly agree	Disagree	Neutral	Agree	Strongly agree
Failure to use diagnostic tests, while the burden and impact of undergoing diagnostic testing for an individual patient and/or the costs of diagnostic testing for society certainly outweigh the burden of setting a diagnosis and the number of patients that can be effectively treated after setting a diagnosis is large.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When patient symptoms persist and diagnostic tests that are considered essential or highly desirable in (inter)national guidelines because these diagnostic tests can confirm or disprove a possible and probable underlying explanation, are not used.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests are omitted that could have yielded relevant findings and would have been effective and socially justifiable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Failure to use diagnostic resources or to request the right diagnostic test at the right time, while the result of such a resource is or may be relevant to the patient or his practitioner for diagnosis, treatment or prognosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests are not used or not sufficiently requested and relevant follow-up of previous deviations is not performed, while there is a clinical question that can potentially be answered.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If too few diagnostic tests have been performed and too much uncertainty remains when setting a diagnosis or guiding a treatment, while a test is available that can reduce that uncertainty.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Results basic principles

	Median	Range	Percentage agreement
Effectiveness	5	4 – 5	100%
Efficiency	4	3 – 5	92%
Relevance	4	3 – 5	92%
Ethical justification	4	3 – 5	67%

Results statements effectiveness, efficiency, relevance, ethical justification

	Median	Range	Percentage agreement
Statements of effectiveness			
The use of diagnostic tests that are necessary for setting a diagnosis as correctly and quickly as possible, avoiding unnecessary and duplicate diagnostic tests whenever possible.	4	1 – 5	67%
Performing medically necessary diagnostic tests on a patient, to set the most likely and accurate diagnosis.	4,5	3 – 5	83%
When the diagnostic resources used, and the results that follow therefrom, are used to guide the diagnosis, treatment, or prognosis of a patient or patient group.	4	1 – 5	50%
The right test at the right time for the patient, in which the requester will adjust his treatment plan based on the results, and that other healthcare providers (at referral) can continue on this basis.	4	2 – 5	75%
The effective use of the concerning diagnostic tests for the correct clinical purposes and clinical question, to achieve the intended effect.	4	2 – 5	58%
The correct diagnostic tests are requested, according to the guidelines, for the intended purpose of specifically answering a diagnostic question.	3	1 – 5	8%
Diagnostic tests of which each step or test is relevant to set a diagnosis or to allow treatment to be targeted and safe.	3,5	2 – 5	50%
Statements of efficiency			
Performing a medically necessary diagnostic test on a patient, at the right time with the to set a diagnosis as quickly as possible, using as few resources as possible and minimizing the burden to the patient.	4,5	3 – 5	92%
The use of a limited number of requests and determinations to answer the referrer's clinical question, to set a diagnosis, and to determine any follow-up policy, where the aim must be that aim to improve the patient's health.	3	2 – 5	42%
Diagnostic tests that are used sensibly and effectively, in which not only the efficiency of the diagnostic method is relevant, but also the use of diagnostic tests provides an efficient and meaningful answer to the concerning issue.	4	1 – 5	67%
The use of diagnostic tests is effective, targeted, cost-efficient, and thus optimal use of effort and resources concerning the expected diagnostic gain.	4	3 – 5	67%



Statements of relevance			
The problem-driven requesting of correct and relevant diagnostic tests resulting in findings that are important for setting a diagnostic conclusion and for the concerning question or decision-making.	4	2 – 5	83%
Diagnostic tests that can differentiate between different diagnoses or support treatment choices that are useful for the diagnosis, health monitoring, treatment, and/or prognosis of the patient.	4	3 – 5	83%
When the correct diagnostic tests are used for the correct (composition) of complaints, not too many and not too few, and with a view to and input from the interests of the patient on which policy can be made that contributes to the diagnosis, treatment, and prognosis.	4	3 – 5	58%
Performing the correct diagnostic tests on a patient, to set a diagnosis, provided that undergoing the tests outweighs setting a diagnosis and subsequently initiating meaningful, targeted, and effective medical treatment.	4	2 – 5	83%
Statements of ethical justification (not included in the third survey)			
Performing sensible diagnostic testing of a patient that does not cause harm to the patient and which has added value for the individual, taking into account his or her life expectancy and life wishes, and taking into account the social health care budget.	5	3 – 5	92%
Diagnostic tests that are performed in an equitable and scientifically justified manner and only used necessary for setting a diagnosis that can contribute to answering the clinical question.	3,5	2 – 5	50%
The added value of diagnostic testing to the possible diagnosis or treatment justifies the possible outcomes of testing and the adverse impact that diagnostic tests may have on the patient or society.	4	2 – 5	58%
That the diagnostic tests used should provide an advantage for the patient and only detect something that leads to serious consequences in a reasonable period, or something that can be addressed to prevent later complications and unfavorable outcomes.	3	2 – 5	42%
The responsible use of certain resources for a diagnostic purpose, whereby anyone who has the right or necessity for a test can obtain it and that this is not compromised by wasting these tests on irrelevant diagnostic tests.	3	2 – 4	33%
The right test, with the right intention at the right time for the patient based on the expertise of the requester with transparency about how the diagnostic tests are performed and about the costs, when diagnostic tests are used for other purposes, the patient should always agree.	4	2 – 5	58%
Results statements correct use and over- and underuse			
	Median	Range	Percentage agreement
Statements of correct use			
If diagnostic tests are effective and ethical, and taking into account a good use of resources and whereby relevant information to set a correct diagnosis is obtained.	4	3 – 5	92%



If exactly those diagnostic tests have been performed that support the correct diagnosis, the monitoring of the patient's health, or the safe and targeted performance of treatment.	4	3 – 5	75%
When the current guidelines or standards have been correctly interpreted and followed, complemented by diagnostic reviewing, and the patient has been involved in the assessment and has been able to make an appropriate contribution to the decision-making process.	3	2 – 4	42%
When there is effective and efficient requesting behavior, where the correct diagnostic test is requested at the right time for the patient.	4	2 – 5	67%
When there is effective, efficient, ethical, and relevant use of diagnostic testing where relevant information is obtained to make policy and where there is no under- or overuse.	5	3 – 5	83%
Statements of overuse			
When the results of a diagnostic test no longer contribute to the diagnosis or treatment policy.	4	3 – 5	83%
When diagnostic tests are performed without any underlying reason or based on improper grounds that do not provide an answer to the patient's health situation and is not necessary for the medical diagnostic process.	5	4 – 5	100%
When diagnostic tests are used when the burden and impact of undergoing diagnostic testing for an individual patient and/or the costs of diagnostic testing for society do not outweigh the added value of setting a diagnosis and the number of patients that can be treated more effectively after setting a diagnosis is limited.	4	2 – 5	67%
Diagnostic tests that are used without the patient's (or practitioner's) questioning or decision-making properly processed and without doing justice to the dilemmas that lie beneath each diagnostic assessment.	4	2 – 4	58%
When diagnostic tests are not effective and/or socially justifiable, because too many resources are used and/or information is obtained that is not entirely relevant for setting a diagnosis or guiding a treatment.	4,5	3 – 5	83%
Statements of underuse			
Failure to use diagnostic tests, while the burden and impact of undergoing diagnostic testing for an individual patient and/or the costs of diagnostic testing for society certainly outweigh the burden of setting a diagnosis and the number of patients that can be effectively treated after setting a diagnosis is large.	5	4 – 5	100%
When patient symptoms persist and diagnostic tests that are considered essential or highly desirable in (inter)national guidelines because these diagnostic tests can confirm or disprove a possible and probable underlying explanation, are not used.	4	3 – 5	75%
When diagnostic tests are omitted that could have yielded relevant findings and would have been effective and socially justifiable.	4,5	2 – 5	83%



Failure to use diagnostic resources or to request the right diagnostic tests at the right time, while the result of such a resource is or may be relevant to the patient or his practitioner for diagnosis, treatment or prognosis.	4	4 – 5	100%
When diagnostic tests are not used or not sufficiently requested and relevant follow-up of previous deviations is not performed, while there is a clinical question that can potentially be answered.	4	3 – 5	58%
If too few diagnostic tests have been performed and too much uncertainty remains when setting a diagnosis or guiding a treatment, while a test is available that can reduce that uncertainty.	4	2 – 5	67%



Appendix V Delphi questionnaire round 3 – 8th June 2020

Dear participant,

Following your participation in the study 'defining of over- and underuse of laboratory diagnostics' you are given the first survey. This study aims to investigate expert opinions to reach a consensus on the definition of over- and underuse.

In round two, 12 of 13 invited participants completed the survey. The results of the second round are that effectiveness, efficiency, and relevance will be included in the candidate definitions of over- and underuse and ethical justification will be excluded.

In round three, the statements, of which at least 65% of the panel members 'agreed' or 'strongly agreed' with are presented again. But this time in the form of smaller partial aspects. Once again, we request you to indicate how important you consider a specific aspect being part of the final definition on a Likert scale of 1 to 5 (1 = strongly disagree - 5 = strongly agree).

As it concerns your opinion, there are no correct or incorrect answers. Should you have any further questions regarding the survey, your participation, or this questionnaire, please contact us via e-mail: xxx or mob: xxx

Q1 Which of the following aspects do you think is best suited to 'effective use of diagnostic tests'? For me, effective use of diagnostic tests means ...

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Setting a correct diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setting a quick diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Avoiding unnecessary diagnostic tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Avoiding duplicate diagnostic tests	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setting the most accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setting the most probable diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The right test at the right time for the patient.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The requester can adjust his treatment plan based on the results.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Based on the results, another healthcare provider can continue on this basis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Performing diagnostic tests that are considered necessary from a medical point-of-view.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

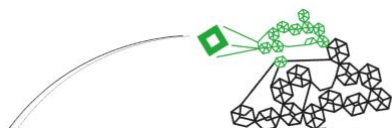


Q2 Which of the following aspects do you think is best suited to 'efficient use of diagnostic tests'? For me, efficient use of diagnostic tests means ...

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Performing diagnostic tests that are considered necessary from a medical point-of-view.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Performing diagnostic tests at the right time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setting a diagnosis using as few resources as possible.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The patient is burdened as little as possible.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests are used sensibly.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that are used effectively.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that provide an answer to the concerning issue.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic resources are used targeted.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic resources are used effectively.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic resources used cost-effectively.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Optimal use of resources concerning the expected diagnostic gain.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Optimal use of effort concerning the expected diagnostic gain.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q3 Which of the following aspects do you think is best suited to 'relevant use of diagnostic tests'? For me, relevant use of diagnostic tests means ...

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
An outcome that is important for setting a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
An outcome that is important for the concerning question.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
An outcome that is important for the decision-making process	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that can differentiate between different diagnoses.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that support treatment choices.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that are useful for diagnosing the patient.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that are useful for the patient's treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that are useful for patient prognosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Diagnostic tests that are useful for the patient's health monitoring.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



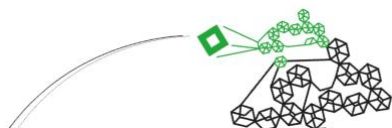
Undergoing the diagnostic tests outweighs setting a diagnosis and subsequently starting treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Problem-driven requesting of diagnostic tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Requesting the correct diagnostic tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Requesting the relevant diagnostic tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q4 Which of the following aspects do you think is best suited to 'correct use of diagnostic tests'? For me, correct use of diagnostic tests means ...

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Effective diagnostic tests.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Obtaining information to set a correct diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use of diagnostic tests that support the correct diagnosis setting.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use of diagnostic tests that support the safe performance of treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use of diagnostic tests that support the targeted performance of treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Use of diagnostic tests that support the monitoring of the patient's health.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Using diagnostic tests in which relevant information is obtained to make policy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Effective request behavior.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Efficient request behavior.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The right test at the right time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q5 Which of the following aspects do you think is best suited to 'overuse of diagnostic tests'? For me, overuse of diagnostic tests means ...

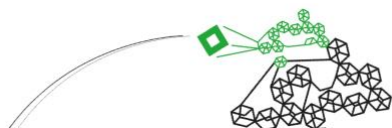
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
When the results of diagnostic tests do not contribute to the diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When the results of diagnostic tests do not contribute to the treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When the results of diagnostic tests do not contribute to the policy.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Performing diagnostic tests without any underlying reason.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Performing diagnostic tests that do not provide an answer to the patient's health situation.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Performing diagnostic tests that are not necessary for the medical diagnostic process.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of diagnostic tests if the burden of undergoing diagnostic testing does not outweigh the added value of making a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of diagnostic tests if the impact of undergoing diagnostic testing does not outweigh the added value of setting a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of diagnostic tests if the costs of undergoing diagnostic testing do not outweigh the added value of making a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The use of diagnostic tests when the number of patients that can be effectively treated after setting a diagnosis is limited.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests are not effective.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests are not socially justifiable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Where information is obtained that is not relevant for making a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Where information is obtained that is not relevant to the treatment.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q6 Which of the following aspects do you think is best suited to 'underuse of diagnostic tests'? For me, underuse of diagnostic tests means ...

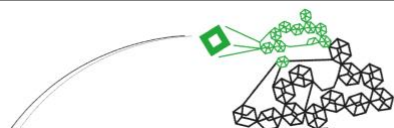
	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Failure to use diagnostic tests while the burden of undergoing diagnostic testing certainly outweighs the burden of setting a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Failure to use diagnostic tests while the impact of undergoing diagnostic testing certainly outweighs the benefits of setting a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Failure to use diagnostic tests while the costs of undergoing diagnostic testing certainly outweigh the costs of setting a diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



Failure to use diagnostic tests while the number of patients that can be effectively treated after setting a diagnosis is large.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests considered essential or highly desirable in (inter)national guidelines, are not used.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If diagnostic tests had been effective.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If diagnostic tests had been socially justifiable.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
When diagnostic tests could have yielded relevant findings.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Failure to use a diagnostic resource or to request the right diagnostic tests at the right time.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
If too much uncertainty remains when setting a diagnosis or guiding a treatment, while a diagnostic test is available that can reduce that uncertainty.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Results of the aspects effectiveness, efficiency, relevance, ethical justification

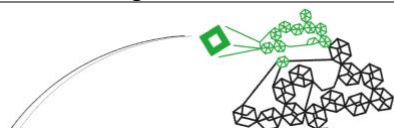
	Median	Range	Percentage agreement
Effectiveness			
Setting a correct diagnosis.	5	4 – 5	100%
Setting a quick diagnosis.	3	3 – 4	42%
Avoiding unnecessary diagnostic tests.	4	3 – 5	92%
Avoiding duplicate diagnostic tests.	4	2 – 5	73%
Setting the most accurate diagnosis.	5	2 – 5	91%
Setting the most probable diagnosis.	4	2 – 5	67%
The right test at the right time for the patient.	4	3 – 5	83%
The requester can adjust his treatment plan based on the results.	4,5	1 – 5	67%
Based on the results, another healthcare provider can continue on with this basis.	4	1 – 5	67%
Performing diagnostic tests that are considered necessary from a medical point-of-view.	4	2 – 5	58%
Efficiency			
Performing diagnostic tests that are considered necessary from a medical point-of-view.	4	2 – 5	67%
Performing diagnostic tests at the right time.	4	3 – 5	75%
Setting a diagnosis using as few resources as possible.	4	2 – 5	83%
The patient is burdened as little as possible.	4	3 – 4	67%
Diagnostic tests used sensibly.	4	2 – 5	67%
Diagnostic tests are used effectively.	4	3 – 5	92%
Diagnostic tests that provide an answer to the concerning issue.	3	2 – 5	42%



Diagnostic test resources are used targeted.	4	2 – 5	58%
Diagnostic test resources are used effectively.	4	3 – 5	83%
Diagnostic resources used cost-effectively.	4	3 – 5	92%
Optimal use of resources concerning the expected diagnostic gain.	4	4 – 5	100%
Optimal use of effort concerning the expected diagnostic gain.	4	3 – 5	83%
Relevance			
An outcome that is important for setting a diagnosis.	5	4 – 5	100%
An outcome that is important for the concerning question.	4	3 – 5	92%
An outcome that is important for the decision-making process.	5	4 – 5	100%
Diagnostic tests that can differentiate between different diagnoses.	4	3 – 5	58%
Diagnostic tests that support treatment choices.	4	3 – 5	83%
Diagnostic tests that are useful for diagnosing the patient.	4	3 – 5	91%
Diagnostic tests that are useful for the patient's treatment.	4	2 – 5	83%
Diagnostic tests that are useful for patient prognosis.	4	2 – 5	58%
Diagnostic tests that are useful for the patient's health monitoring.	4	2 – 5	67%
Undergoing the test outweighs setting a diagnosis and subsequently starting treatment.	3	2 – 5	42%
Problem-driven requesting of diagnostic tests.	3,5	3 – 5	50%
Requesting the correct diagnostic tests.	4	3 – 5	83%
Requesting the relevant diagnostic tests.	4	3 – 5	75%

Results of the aspects correct use and over- and underuse

	Median	Range	Percentage agreement
Aspects of correct use			
Effective diagnostic tests.	4	3 – 5	69%
Obtaining information to set a correct diagnosis.	4	2 – 5	69%
Use of diagnostic tests that support the correct diagnosis.	4	3 – 5	75%
Use of diagnostic tests that support the safe performance of treatment.	4	2 – 5	54%
Use of diagnostic tests that support the targeted performance of treatment.	4	3 – 5	69%
Use of diagnostic tests that support the monitoring of the patient's health.	4	3 – 5	54%
Using diagnostic tests in which relevant information is obtained to make policy.	4	2 – 5	85%
Effective request behavior.	4	3 – 5	77%
Efficient request behavior.	4	2 – 5	62%
The right test at the right time.	4	3 – 5	92%
Aspects of overuse			
When the results of diagnostic tests do not contribute to the diagnosis.	4	2 – 5	69%



When the results of diagnostic tests do not contribute to the treatment.	4	2 – 5	77%
When the results of diagnostic tests do not contribute to the policy.	4	2 – 5	62%
Performing diagnostic tests without any underlying reason.	5	4 – 5	100%
Performing diagnostic tests that do not provide an answer to the patient's health situation.	4	3 – 5	85%
Performing diagnostic tests that are not necessary for the medical diagnostic process.	4	1 – 5	85%
The use of diagnostic tests if the burden of undergoing diagnostic testing does not outweigh the added value of setting a diagnosis.	4	1 – 5	69%
The use of diagnostic tests if the impact of undergoing diagnostic testing does not outweigh the added value of setting a diagnosis.	4	1 – 5	69%
The use of diagnostic tests if the costs of undergoing diagnostic testing do not outweigh the added value of setting a diagnosis.	4	1 – 4	69%
The use of diagnostic tests when the number of patients that can be effectively treated after setting a diagnosis is limited.	3	2 – 5	15%
When diagnostic tests are not effective.	3	2 – 5	46%
When diagnostic tests are not socially justifiable.	3	3 – 5	23%
Where information is obtained that is not relevant for setting a diagnosis.	4	2 – 5	54%
Where information is obtained that is not relevant to the treatment.	4	2 – 5	62%
Aspects of underuse			
Failure to use diagnostic tests while the burden of undergoing diagnostic testing certainly outweighs the burden of setting a diagnosis.	4	2 – 5	77%
Failure to use diagnostic tests while the impact of undergoing diagnostic testing certainly outweighs the benefits of setting a diagnosis.	4	2 – 5	77%
Failure to use diagnostic tests while the costs of undergoing diagnostic testing certainly outweigh the costs of setting a diagnosis.	4	2 – 5	77%
Failure to use diagnostic tests while the number of patients that can be effectively treated after setting a diagnosis is large.	4	3 – 5	77%
When diagnostic tests considered essential or highly desirable in (inter)national guidelines, are not used.	4	3 – 5	92%
If diagnostic tests had been effective.	4	3 – 5	62%
If diagnostic tests had been socially justifiable.	3	2 – 4	23%
When diagnostic tests could have yielded relevant findings.	4	2 – 5	92%
Failure to use a diagnostic resource or to request the right diagnostic tests at the right time.	4	3 – 5	85%
If too much uncertainty remains when setting a diagnosis or guiding a treatment, while a diagnostic test is available that can reduce that uncertainty.	4	3 – 5	62%



Appendix VI Delphi questionnaire round 4 – 2nd July 2020

Dear participant,

Following your participation in the study 'defining of over- and underuse of laboratory diagnostic tests' you are given the fourth and last survey. This study aims to investigate expert opinions to reach a consensus on the definition of over- and underuse.

In this last survey, a total of 28 candidate definitions are presented, of which 14 representing overuse and 14 representing underuse. The definitions are composed of one or more aspects, more aspects are repeatedly added, which makes the definition more extensive. The definitions are therefore similar to each other, but each time slightly different. We ask you to rate these candidate definitions on a Likert scale from 1 to 10 (1 = extremely disagree - 10 = excellent).

In concluding, we ask you once again to give your consent (afterward) to participate in this study. We ask this because there may be respondents who have not yet given consent because they did not participate in the first survey.

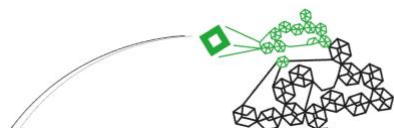
As it concerns your opinion, there are no correct or incorrect answers. Should you have any further questions regarding the survey, your participation, or this questionnaire, please contact us via e-mail: xxx or mob: xxx

Q2 Overuse is the use of diagnostic tests without any underlying reason or when the outcome is not relevant for...:

	Extremely disagree	Strongly disagree	Disagree	Slightly disagree	Neutral	Slightly agree	Agree	Strongly agree	Extremely agree	Excellent
The clinical question.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical decision-making.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setting a correct and accurate diagnosis.										
The clinical question, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical decision-making, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



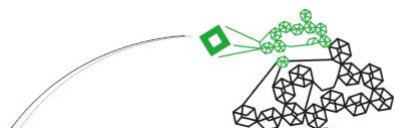
Setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical question or decision-making.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical question, or for setting a correct and accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical decision-making, or for setting a correct and accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical question or decision-making, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical question, or for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical decision-making, or for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical question or decision-making, or for setting a correct and accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The clinical question or decision-making, or for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient,	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



or cost-effective when considered against the expected health benefits.

Q1 Underuse is the failure to use diagnostic tests when this could have yielded relevant findings for...:

	Extremely disagree	Strongly disagree	Disagree	Slightly disagree	Neutral	Slightly agree	Agree	Strongly agree	Extremely agree	Excellent
The concerning question.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The decision-making process.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setting a correct and accurate diagnosis.										
The concerning question, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The decision-making process, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Setting a correct and accurate diagnosis, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The concerning question, or the decision-making process.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The concerning question, or for setting a correct and accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The decision-making process, or for setting a correct and accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The concerning question, or the decision-making process, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The concerning question, or for setting a correct and accurate diagnosis or	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The decision-making process, or for setting a correct and accurate diagnosis, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The concerning question, or the decision-making process, or for setting a correct and accurate diagnosis.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
The concerning question, or the decision-making process, or for setting a correct and accurate diagnosis or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q3 I have read and understood the information letter and consent form and consent to everything contained herein.

- I give consent
- I do not give consent

Results candidate definitions overuse (overuse is the use of diagnostic tests without any underlying reason or when the outcome is not relevant for...:)

	Median	Range	Total sum of ratings	Percentage agreement
The clinical question.	7	6-9	80	36%
The clinical decision-making.	7	6-9	83	45%
Setting a correct and accurate diagnosis.	7	5-9	80	36%
The clinical question, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	7	6-9	78	18%
The clinical decision-making, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	7	6-9	80	36%
Setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	7	6-9	80	36%
The clinical question or decision-making.	7	5-9	81	36%



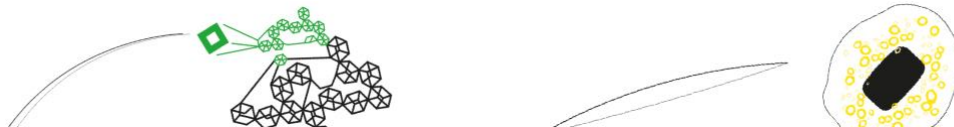
The clinical question, or for setting a correct and accurate diagnosis.	7	6-9	78	18%
The clinical decision-making, or for setting a correct and accurate diagnosis.	7	6-9	80	36%
The clinical question or decision-making, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	7	6-9	79	27%
The clinical question, or for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	7	6-9	80	36%
The clinical decision-making, or for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	7	5-9	80	36%
The clinical question or decision-making or for setting a correct and accurate diagnosis.	7	6-9	84	45%
The clinical question or decision-making, for setting a correct and accurate diagnosis, or when the use of diagnostic tests is not optimal, efficient, or cost-effective when considered against the expected health benefits.	8	6-10	86	55%

Results candidate definitions underuse (underuse is the use of diagnostic tests without any underlying reason or when the outcome is not relevant for...:)

	Median	Range	Total sum of ratings	Percentage agreement
The concerning question.	7	2-10	71	27%
The decision-making process.	8	7-10	87	55%
Setting a correct and accurate diagnosis.	8	6-9	84	64%
The concerning question, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	7	4-8	70	9%
The decision-making process, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	7	6-8	76	18%
Setting a correct and accurate diagnosis or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	7	5-9	77	27%
The concerning question, or the decision-making process.	7	5-9	84	45%
The concerning question, or for setting a correct and accurate diagnosis.	7	5-9	78	36%
The decision-making process, or for setting a correct and accurate diagnosis.	8	5-9	85	64%
The concerning question, or the decision-making process, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	7	6-9	79	36%
The concerning question, or for setting a correct and accurate diagnosis or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	7	6-9	77	36%

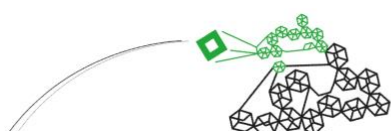


The decision-making process, or for setting a correct and accurate diagnosis, or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	7	6-9	80	45%
The concerning question or decision-making or for setting a correct and accurate diagnosis.	7	6-9	80	77
The concerning question, the decision-making process or for setting a correct and accurate diagnosis or when diagnostic tests are considered essential or highly desirable in (inter)national guidelines.	7	6-8	77	36%



References

1. Lippi GaBCaCM. Inappropriateness in laboratory medicine: An elephant in the room? *Annals of Translational Medicine*. 2017;5.
2. Plebani M. Appropriateness in programs for continuous quality improvement in clinical laboratories. *Clinica Chimica Acta*. 2003;333(2):131-9.
3. Cadamuro JaIMaCMaNMaH-BEavMAaLGA-SA-M. Managing inappropriate utilization of laboratory resources. *Diagnosis*. 2018.
4. Gion MaTCaFA. Appropriateness of tumor marker request: A case of study. *Annals of Translational Medicine*. 2017;5:274-.
5. Bogovac-Stanojevic N, Jelic-Ivanovic Z. The Cost-effective Laboratory: Implementation of Economic Evaluation of Laboratory Testing. *J Med Biochem*. 2017;36(3):238-42.
6. Meidani Z, Farzandipour M, Farrokhan A, Haghighat M. A review on laboratory tests' utilization: A trigger for cutting costs and quality improvement in health care settings. *Med J Islam Repub Iran*. 2016;30:365.
7. Cappelletti P. Appropriateness of diagnostics tests. *International Journal of Laboratory Hematology*. 2016;38:91-9.
8. Duddy C, Wong G. Explaining variations in test ordering in primary care: protocol for a realist review. *BMJ Open*. 2018;8(9):e023117.
9. Hauser RG, Shirts BH. Do We Now Know What Inappropriate Laboratory Utilization Is?: An Expanded Systematic Review of Laboratory Clinical Audits. *American Journal of Clinical Pathology*. 2014;141(6):774-83.
10. Khan KS, Kunz R, Kleijnen J, Antes G. Five steps to conducting a systematic review. *J R Soc Med*. 2003;96(3):118-21.
11. Hasson FaKSaMH. Research guidelines for the Delphi Survey Technique. *Journal of advanced nursing*. 2000;32:1008-15.
12. Zhi M, Ding EL, Theisen-Toupal J, Whelan J, Arnaout R. The landscape of inappropriate laboratory testing: a 15-year meta-analysis. *PLoS One*. 2013;8(11):e78962.
13. Vrijens BEL, Naaktgeboren CA, Vos LM, van Solinge WW, Kaasjager HAH, ten Berg MJ. Inappropriate laboratory testing in internal medicine inpatients: Prevalence, causes and interventions. *Annals of Medicine and Surgery*. 2020;51:48-53.
14. O'Sullivan JaAAaNBaPRaAJaRNaHC. Overtesting and undertesting in primary care: a systematic review and meta-analysis. *BMJ Open*. 2018;8:e018557.
15. Chami N, Simons JE, Sweetman A, Don-Wauchope AC. Rates of inappropriate laboratory test utilization in Ontario. *Clinical Biochemistry*. 2017;50(15):822-7.
16. Ferraro S, Panteghini M. The role of laboratory in ensuring appropriate test requests. *Clinical Biochemistry*. 2017;50(10):555-61.
17. Lippi GaMC. The biomarker paradigm: Between diagnostic efficiency and clinical efficacy. *Polskie Archiwum Medycyny Wewnetrznej*. 2015;125.
18. Mrazek CaSA-MaSMavMAaCMABJMaNMaLGAH-BEa. Inappropriate use of laboratory tests: How availability triggers demand – examples across Europe. *Clinica chimica acta; international journal of clinical chemistry*. 2020.
19. Miyakis S, Karamanof G, Lontos M, Mountokalakis TD. Factors contributing to inappropriate ordering of tests in an academic medical department and the effect of an educational feedback strategy. *Postgrad Med J*. 2006;82(974):823-9.
20. Sarkar Mayukh K, Botz Chad M, Laposata M. An assessment of overutilization and underutilization of laboratory tests by expert physicians in the evaluation of patients for bleeding and thrombotic disorders in clinical context and in real time. *Diagnosis* 2017. p. 21.
21. Bindraban RS, Ten Berg MJ, Naaktgeboren CA, Kramer MHH, Van Solinge WW, Nanayakkara PWB. Reducing Test Utilization in Hospital Settings: A Narrative Review. *Ann Lab Med*. 2018;38(5):402-12.
22. Cadamuro J, Gaksch M, Wiedemann H, Lippi G, von Meyer A, Pertersmann A, et al. Are laboratory tests always needed? Frequency and causes of laboratory overuse in a hospital setting. *Clinical Biochemistry*. 2018;54:85-91.
23. Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. *Journal of Advanced Nursing*. 2006;53(2):205-12.



24. Diamond IR, Grant RC, Feldman BM, Pencharz PB, Ling SC, Moore AM, et al. Defining consensus: A systematic review recommends methodologic criteria for reporting of Delphi studies. *Journal of Clinical Epidemiology*. 2014;67(4):401-9.
25. Keeney S, Hasson F, McKenna HP. A critical review of the Delphi technique as a research methodology for nursing. *International Journal of Nursing Studies*. 2001;38(2):195-200.
26. Holey EA, Feeley JL, Dixon J, Whittaker VJ. An exploration of the use of simple statistics to measure consensus and stability in Delphi studies. *BMC Med Res Methodol*. 2007;7:52.

