Exploring Perspectives and Preferences for Over-Diagnosis and Overtreatment of Thyroid Nodules using a MCDA Method

Linda Karrer
Friedrich-Alexander-Universitat Erlangen-Nurnberg

Shixuan Zhang (✉️ shixuan.zhang@fau.de)
Friedrich-Alexander-Universitat Erlangen-Nurnberg
https://orcid.org/0000-0002-8427-4943

Thomas Kühlein
Friedrich-Alexander-Universitat Erlangen-Nurnberg

Peter L Kolominsky-Rabas
Friedrich-Alexander-Universitat Erlangen-Nurnberg

Research

Keywords: Thyroid nodules, Multi criteria decision Analysis, MCDA, EVIDEM, Overdiagnosis, Adverse Cascade Effects, Health Technology Assessment (HTA)

DOI: https://doi.org/10.21203/rs.3.rs-89748/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

**Background:** The detection of thyroid cancer has rapidly increased over the last few decades without an increase in disease-specific mortality. Several studies claim that the diagnosis of thyroid nodules through routine ultrasound imaging is often the trigger for cascade effects leading to unnecessary follow-up over many years or to invasive treatment. The objective of this study was to explore physicians’ and patients’ insights and preferences for diagnosis and treatment of thyroid nodules as well as awareness of over-diagnosis and overtreatment.

**Methods:** An online survey was developed using a comprehensive multi criteria decision Analysis (MCDA) framework, the EVIdence based Decision-Making (EVIDEM). The EVIDEM core model used in this study encompassed 13 quantitative criteria and four qualitative criteria. Participants were asked to provide weights for each criterion, performance scores for appraising the thyroid nodules interventions and their consideration of impact of contextual criteria. Normalized weights and standardized scores were combined to calculate a value contribution across all participants, additionally differences across participants were explored.

**Results:** Of all 105 participants, there are 48 patients, 31 physicians and 26 normal citizens. The highest value estimate of the intervention on thyroid nodules reached 0.401 for citizens’ group and lowest value of 0.287 was reported for physicians’ group, compared to 0.359 for all participants. The highest value contributors were efficacy and prevention. Compared to all participants, physicians gave less weight to quality of life and prevention benefits. In addition, physicians’ group gave less score on the treatment such like thyroid surgery compared to other two groups. 59% participants considered healthcare system had a positive impact on the diagnostic of thyroid nodules.

**Conclusion:** Our study shows participants’ preferences on each criterion, i.e., physician indicated keeping the intervention safe and effective more important, patients indicated quality of life after receiving intervention more important. Through comparison among participants, differences have been highlighted, which can make better communication between physicians and patients. This study provided a supportive decision making for health providers when they conduct researches on thyroid nodules.

**Trial registration:** This study did not report the results of a health care intervention on human participants.

1. **Background**

Thyroid nodules (TN) refers to an abnormal growth of thyroid cells that form a lump within the thyroid gland [1]. They are mostly benign and harmless, but also can be malignant or lead to hyperthyroidism. The nature of most TN is not clear until by imaging alone. TN are very common, according to the American Thyroid Association, about 50% of all people by age 60 have TN, over 90% of them are noncancerous [1]. In order to detect thyroid cancer, most TN if seen in ultrasound need further evaluation, such as control thyroid ultrasonography over time, thyroid hormone test, scintigraphy or fine needle aspiration cytology. The detection and thus the prevalence of thyroid cancer has rapidly increased over the last few decades.
without an increase in disease specific mortality [2]. Thyroid surgery is more common in Germany in contrast with the situation in some other countries, in a 2015 study, the rate was about four times higher than that in England and seven times higher than the rate of Netherland [3]. Especially in Germany the diagnosis of TN leads to a growing amount of thyroid surgery after which histology reveals benign nodules that would not have needed to be removed [4]. Several studies from industrialized countries have shown the increasing numbers of thyroid cancer with constant or decreasing mortality [2]. Significant increase in TN and thyroid cancer in developed countries is attributed mainly to an unapproved diagnostic imaging of the thyroid gland by ultrasound [5, 6]. We claim that the diagnose of TN through routine ultrasound imaging is often the trigger for cascade effects leading to unnecessary follow-up over many years or to invasive treatment [7].

“PRO PRICARE (Preventing Overdiagnosis in Primary Care)” is a network of health services research at the Friedrich Alexander- Universität Erlangen-Nürnberg (FAU) funded by the Germany Federal Ministry for Education and Research (BMBF) [8]. The project was led by the Institute of General Practice, in cooperation with other institutes at the FAU and the University Hospital Erlangen. PRO-PRICARE was divided into three sub-projects, one of them being “Adverse cascade effects” (ACE). Cascade effects are defined as “processes that, once initiated, will stepwise proceed until their seemingly inevitable results”. One common example is the treatment of patients with TN being potentially malignant discussed above. The aim of the ACE was to learn more about the existence and reasons for cascades in the care for patients with TN. One of the studies within ACE is the one here reported. Although overdiagnosis and overtreatment in the care of patients with TN in Germany are well described, little is known about the perspectives and preferences of physicians and patients.

The objective of this study was to explore physicians’ and patients’ insights and preferences for diagnosis and treatment of TN as well as awareness of over-diagnosis and overtreatment, using a multi criteria decision analysis (MCDA) method.

2. Methods

2.1 Study design

The EvIdence based decision-making (EVIDEM, 10th Edition 2019) framework was selected to investigate participants’ insights and preferences. EVIDEM framework is designed to reflect and to stimulate structured reflection and pragmatic collection of preferences on healthcare interventions from all participants, through a broad spectrum of quantitative and qualitative criteria [9]. In this study, we provided synthesized data based on analysis of the intervention on TN for each of these criteria to create a TN specific EVIDEM core model online questionnaire. Participants were investigated to fill in an online questionnaire to collect their insights and preferences. Participants included physicians, patients who ever had thyroid diseases and normal citizens who had not thyroid diseases before, named as citizens in this study, stemming from treating or living with the disease, or having great interest on this topic.
Participants were recruited through public access like medical networks, regional distributors and newspaper announcement. All participants were recruited anonymously and voluntarily.

2.2 Online questionnaire design and conduct

An online survey was created in line with the core model of EVIDEM framework for the participants. The EVIDEM core model used in this study encompassed five categories in total 13 quantitative criteria, while the contextual tool consisted of four qualitative criteria. The description of each category and criterion used in this study is shown in Table 1. A definition of all criteria as well as background knowledge such as sociodemographic data used in this survey was provided to participants in the online questionnaire in German language. To provide sufficient evidence to appraise each criterion, a literature review was used to obtain relevant information on TN and its current management [8, 10, 11]. All participants were invited to fill in an online questionnaire. The online questionnaire was constructed through the survey software EFS Survey by UNIPARK (https://www.unipark.com/umfragesoftware/) [12]. Informed written consent and online questionnaire were approved by „Data protection officer“ at FAU according to the Bavarian State Representative Data Protection (https://www.datenschutz-bayern.de/vorstell/impressum.html). We contacted an online survey open to the public, Informed written consent was clicked while fulfilling the online survey by each participant. Participants were assured that the research would not contain their personal identifying information.

In the first part of the survey, participants’ perspectives on what matters most in general, i.e., which criteria contribute the most to the value of healthcare interventions, was captured by weight. Our study used a 5-point weighting scale (1 = lowest relative importance, 5 = highest relative importance). In the second part of the survey, participants were asked to appraise the actual intervention on TN about its performance for each criterion, which captured by score. Participants scored performance of the intervention on TN using two types of scoring scale, for non-comparative criteria from 0 to + 5, for comparative criteria from − 5 to + 5. Higher score indicates better performance. The third part of the survey was about qualitative contextual criteria, participants indicated whether consideration of a given criterion had a negative, neutral, or positive impact on the decision about the intervention on TN.
Table 1
MCDA EVIDEM Framework 10th Edition adapted to evaluation of intervention on thyroid nodules

| Quantitative Criteria : MCDA Core Model |
|-----------------------------------------|
| **Category**                             | **Criterion**                      | **Short terms showed in Figures** |
| Disease needs                           | Disease severity                   | Disease severity                  |
|                                         | Size of affected population        | Affected population               |
|                                         | Unmet needs                        | Unmet needs                       |
| Comparative outcomes of intervention    | Comparative efficacy/effectiveness  | Effectiveness                      |
|                                         | Comparative safety/tolerability     | Safety                            |
|                                         | Comparative patient-perceived health| Quality of life (QoL)             |
| Type of benefit provided by intervention| Type of preventive benefit         | Prevention                         |
|                                         | Type of therapeutic benefit         | Therapy                            |
| Comparative economic consequences of    | Comparative cost of intervention   | Intervention cost                  |
| intervention                           | Comparative other medical costs     | Other medical cost                 |
|                                         | Comparative other non-medical costs | Non-medical cost                   |
| Knowledge about the intervention        | Quality of evidence                 | Evidence                           |
|                                         | Clinical practice guidelines        | Guideline                          |

| Contextual Criteria: MCDA Contextual Tool |
|-------------------------------------------|
| Mandate and scope of the healthcare system| Healthcare system                   |
| System capacity and appropriate use of intervention| System capacity                   |
| Population priorities and access          | Access                               |
| Political/historical/cultural context     | Political context                   |

2.3 Data analysis

Numerical outputs were calculated for each participant. In this study, we used Excel to calculate and analyze weight, score and contextual impact. Mean and standard deviations (SD) were calculated in Excel to quantify the variability as descriptive statistics. Normalized weights were summed up to 1.0 (W_x, \( \Sigma W_x = 1 \)). For example, for one single participant, the normalized weight of each criterion equals to the
weight of this criterion given by this participant divided by the total weights given by this participant. For example, participant A weighted a criterion with a “5”, and total weight of all criteria given by participant A was “50”, then the normalized weight for the criterion given by participant A was 5/50 = 0.1. The value contribution (VCx) of each criterion was calculated following a linear additive model as sum of the products of the normalized weights and standardized scores (= score/5). For example, if a criterion received a normalized weight of 0.05, and a score of 5, its value contribution is 0.05*(5/5) = 0.05. For the evaluation of the contextual criteria, a numerical scale (-1, 0, 1) was used to represent negative impact, no impact and positive impact.

3. Results

The online survey was conducted from 2018 November 20th to 2019 June 30th. At the point of closing the online system on 2019 June 30th, we had received valid data from 105 participants in total. Of all these participants, there are 31 physicians, 48 patients who have or ever had thyroid diseases, and 26 normal citizens who had no thyroid diseases, we called citizens in this study (See Table 2).
Table 2
Sociodemographic data of participants

| Characteristic            | Frequency | Percentage |
|---------------------------|-----------|------------|
| Gender                    |           |            |
| Male                      | 55        | 52.4%      |
| Female                    | 50        | 47.6%      |
| Age (years)               |           |            |
| < 20                      | 0         | 0          |
| 20–29                     | 2         | 1.9%       |
| 30–39                     | 5         | 4.8%       |
| 40–49                     | 15        | 14.3%      |
| 50–59                     | 20        | 19%        |
| 60–69                     | 28        | 26.7%      |
| > 70                      | 35        | 33.3%      |
| Stakeholders              |           |            |
| Physician                 | 31        | 29.5%      |
| Patient                   | 48        | 45.7%      |
| Neither physician nor patient | 26      | 24.8%      |
| Medical specialists of Physicians |   |          |
| General practice          | 18        | 58.1%      |
| Internal medicine         | 6         | 19.4%      |
| Surgery                   | 3         | 9.6%       |
| Radiology                 | 1         | 3.2%       |
| Otolaryngology            | 1         | 3.2%       |
| Other fields              | 2         | 6.4%       |

### 3.1 Perspectives of participants on decision criteria

Regarding weights provided by participants, Fig. 1 shows that the most important criteria was “Comparative effectiveness (0.088 ± 0.010)”, followed by “Type of therapeutic benefit (0.086 ± 0.010)”. 
and then “Disease severity (0.086 ± 0.011)”. Normalized weight across criteria was summing up to 1.0, higher weight indicates that the criterion is more important from participants’ view. As Fig. 1 shows, the least important criteria were three cost consequences of intervention relative criteria. The largest variations in weights were observed for “Cost of non-medical intervention (SD, 0.022)”, “Size of population (SD, 0.021)” and “Unmet needs (SD, 0.020)”. The smallest variations were “Quality of life (SD, 0.008)”, “Comparative effectiveness (SD, 0.010)”, “Type of therapeutic benefit (SD, 0.010)”. Mean normalized weights were also calculated assigned to each criterion by three groups as physicians’ group, patients’ group and citizens’ group (Fig. 2). Physicians weighted “Safety of intervention” higher than the other two groups, for the physicians’ group 0.086 ± 0.011, for patients’ group 0.079 ± 0.014 and citizens’ group 0.079 ± 0.017. Large variance was also showed in ‘Type of preventive benefits’, for citizens’ group was 0.082 ± 0.011, for patients’ group was 0.084 ± 0.012, however, in the physicians’ group, the weight was much lower than in the other two groups with a weight of 0.072 ± 0.018. Similar situation also happened to the criterion of “Comparative patient perceived health”, the weights given by physicians 0.074 ± 0.021 lower than the other two groups, for patients 0.085 ± 0.008 and for citizens’ group 0.089 ± 0.009.

### 3.2 Scores

As Fig. 3 shows, for non-comparative criteria, “Type of preventive benefit” received the highest score (0.682 ± 0.261), which shows most participants gave highest performances score on the preventive methods for thyroid disease. Followed by “Type of therapeutic benefit”, this was scored 0.636 ± 0.188. Especially this criterion had the smallest SD, which indicates most participants have an agreement on therapeutic methods are highly useful for treating thyroid disease. For the category “Need of intervention”, “Unmet needs” received the smallest score 0.408 ± 0.251, followed by “Disease Severity”, with a score of 0.488 ± 0.211. But “Size of affected population” in the same category received a higher score 0.533 ± 0.196. For comparative criteria, “Comparative effectiveness” received the highest score, 0.634 ± 0.402. However, in the meantime, the SD is also very high, which indicates some participants gave a very low score. This was also the case to other comparative criteria, all five comparative criteria got relatively higher SD compared to non-comparative criteria. “Comparative safety” in the same category “Treatment Interventions” got a low score of 0.299 ± 0.446. We designed this question as “How do you assess the safety of surgical treatments of TN?” In this category, the criterion “Comparative patient perceived health” got a very low score 0.006 ± 0.435. The question is: “How do you assess the influence of the diagnosis of TN on the quality of life of those affected?”

As shown in Fig. 4, we also calculated mean (SD) standardized scores by different stakeholders as physicians’ group, patients’ group and citizens’ group. The physicians’ group viewed the impact on ‘comparative effectiveness’ and ‘Quality of evidence’ higher than other criteria, however, the patients’ group and the citizens’ group gave ‘Type of preventive benefit’ and ‘Type of therapeutic benefit’ highest scores, respectively. Three criteria that also need to be highlighted are ‘Disease severity’, ‘Comparative
safety’ and ‘Clinical practice guidelines’. Referring to ‘Disease severity’, the score of physicians’ group (0.419 ± 0.209) was smaller than in the patients’ group (0.479 ± 0.192), and much smaller than in the citizens’ group (0.585 ± 0.219). Referring to ‘Comparative safety’, the score of the physicians’ group (0.174 ± 0.534) was smaller than patients’ group (0.388 ± 0.432) and citizens’ group (0.285 ± 0.319). Another interesting criterion was ‘Clinical practice guidelines’, the score of the physicians’ group (0.420 ± 0.215) was also smaller than the other two groups (patients 0.583 ± 0.206, citizens 0.569 ± 0.185).

3.3 Value contribution

Figure 5 shows the mean (SD) value contribution of intervention on TN from all participants after adjusting the performance scores to the weights for each criterion. The mean value estimate of all criteria by all participants was 0.359 on a scale of 0 to 1. The highest value contributors were the ‘Comparative effectiveness’ (0.056 ± 0.004) and the ‘Type of preventive benefits’ (0.056 ± 0.025). The criteria of the ‘Comparative cost consequences-other medical costs’ (-0.018 ± 0.027), the ‘Comparative cost consequences-cost of intervention’ (-0.012 ± 0.028) and the ‘Comparative patients perceived health’ (-0.001 ± 0.037) were negatively contributed to the value.

As shown in Fig. 6, the highest value estimate of the intervention on TN reached 0.401 for citizens’ group and lowest value of 0.287 was reported for physicians’ group. The patients’ group reported a value of 0.368. This figure shows very clear that different stakeholders’ preferences and thoughts are different. In the physicians’ group, the highest value contributors were the ‘Comparative effectiveness’ (0.058 ± 0.042) and the ‘Quality of evidence’ (0.048 ± 0.022). In the patients’ group, the highest value contributors were the ‘Type of preventive benefits’ (0.059 ± 0.022) and the ‘Comparative effectiveness’ (0.058 ± 0.033). In the citizens’ group, the highest value contributors were the ‘Type of preventive benefits’ (0.066 ± 0.021) and the ‘Type of therapeutic benefits’ (0.061 ± 0.019).

3.4 Impacts of contextual criteria

Figure 7 illustrates qualitative contextual criteria. 59% participants considered “Mandate and scope of the healthcare system” (“How do you rate the impact of healthcare system on thyroid diagnostics”) had a positive impact. Consideration of “System capacity and appropriate use of intervention”, 63% participants thought it had a negative impact, and for “Population priorities and access”, with the question “How do you rate the influence of stakeholders on the procedure in thyroid diagnostics”, also 47% participants considered it had a negative impact. Nearly 58% participants thought “Political/historical/cultural context” had no impact on the intervention of thyroid nodules. The overall negative impact, no impact and positive impact was 38%, 35%, and 27%, respectively.

4. Discussion

4.1 Using online platform to conduct MCDA

This study refers to the current issues which are discussed in many countries’ healthcare system [10, 11, 13–16]. The incidence of thyroid cancer has been increasing faster than other cancer because of the
This increased incidence of thyroid cancer with attribution to over-diagnosis has been described in most developed countries where patients have high access to health detection. The study estimated that 70–90% of these patients had asymptomatic lesions during lifetime if ultrasound and other imaging studies were not available [14]. Our study used MCDA to identify the perspective and preference of TN interventions from physicians’, patients’ and citizens’ groups under German healthcare system. EVIDEM was the decision support tool we selected to fulfill this MCDA study.

EVIDEM provides a set of generic decision criteria which were made and selected with the goal to support the substantive legitimacy of the decision with regard to the common goal of healthcare system [9]. To highlight the benefits of using EVIDEM: one is to get preferences of which criteria contribute the most to the worth of healthcare interventions in general, which captured as “weights”; the other is to explore the value of the intervention on TN, which captured by performance scores. Since EVIDEM provides a set of standardized criteria, this gives the chance for researchers to compare different perspectives of the same criterion on different disease. There are more and more researchers using EVIDEM to explore perspectives on healthcare intervention. Most of them recruited a group of people to make a focus group to do interview face to face [17–20]. Our study used an on-line platform to involve more participants [21]. Recruitment of more participants can avoid being influenced by others, such as people in a focus group.

4.2 Perspectives of participants in general on decision criteria

Weights reflect participants’ values and preferences, like what matters most to them. Participants in this study indicated that the most important criteria were ‘Comparative effectiveness’, ‘Type of therapeutic benefit’, ‘Disease severity’, ‘Comparative patients’ perceived health’, ‘Comparative safety’ and ‘Unmet needs’. These results are similar to the results of a survey for the topic of chronic heart disease of all types of stakeholders across the healthcare decision continuum in Germany, which indicated the most important criteria were ‘Clinical effectiveness’, ‘Patients’ perceived health’, ‘Disease severity’, ‘Clinical safety’ and ‘Quality of evidence’ [21].

Compared to physicians, patients tended to assign higher weights to the criteria of patients’ perceived health, higher quality of life for them is much more important after the intervention. This showed also in other study, that patients’ group assigned greater weight to the impact on Health Related Quality of Life [19, 20]. Meanwhile compared to patients, physicians took ‘safety’ more into account. Since physicians are healthcare providers and patients are the receivers of the intervention, physician knows more about what the risk of the intervention, physician indicated keeping the intervention safe and effective were much more important. This difference truly highlights the need for effective communication between physicians and patients, which helps patients to express their need and incorporate their individual priorities in patients-centered healthcare system. Referring to ‘Economic consequence of intervention’, all participants assigned lower values than other criteria, this showed the same in another MCDA study under German healthcare system [21]. These results are in agreement with health professionals, they
wish to help patients without focusing on economic constrains [21, 22]. For the patients, they do not have economic constrains in the context of German healthcare insurance system for most diseases [23].

**4.3 Appraisal of the intervention on thyroid nodules**

Participants used scores to express their views on how each criterion favored for the health intervention based on their own experience and knowledge as well as provided information. Compared to physicians’, patients’ group and citizens’ group highly believed that TN can be prevented and treated. Although this disease is very common, the causes and risk factors of most TN are not clear, it is difficult to prevent this disease [24]. There also one study shows that nutrition can prevent TN, like Selenium [25]. More evidence and other preventive methods we still need more scientific researches.

Referring to ‘Disease severity’, physicians thought TN as a disease with lower severity than the other two groups. This result is in agreement of many studies, although the prognosis is increasing since the beginning of the twenty-first century, the incidence rates of thyroid cancer and mortality have stabilized in more recent years [26]. Referring to ‘Comparative safety’, the score of physicians’ group was smaller than the other two groups, which shows that physicians thought an operation on TN was not as safe as patients and others thought. This result is in agreement with other medical researchers, operative complications can be significant, including permanent hypoparathyroidism, vocal fold paralysis, and airway compromise [27]. For ‘Clinical practice guidelines’, the score of physicians’ group was smaller than the other two groups. As health practitioners, physicians thought they did not follow the recommendation for surgical treatment of a TN in accordance with the guidelines as the others. However other study already proved with the 2009 American Thyroid Association (ATA) guideline revisions, the rapid increase in thyroid cancer incidence rates has recently slowed, especially among small-sized cancers and women [27].

**4.4 Impact of contextual criteria**

48% participants considered “Population priorities and access” having a negative impact on the intervention of TN. This agrees with many studies, that increased incidence of thyroid cancer with attribution to over-diagnosis has been described in most developed countries where patients have high access to health detection. Additional such diagnosis brings a growing amount of thyroid surgery on benign nodules that would not have needed to be removed. Nearly 58% participants thought “Political/historical/cultural context” had no impact on the intervention of TN. This result is in agreement with other studies, possible changes in exposure to risk factors such as diagnostic radiation overweight, diabetes may increase patients’ medical surveillance, and changes in access to health inspection of thyroid gland may also be the likely explanations [10].

**4.5 Limitation of Study**

There are several limitations in this study. Lacking comparators is the second limitation of our study, since the design of EVIDEM framework divided into comparative criteria and non-comparative criteria, analysis in different participants’ groups is our compensation method. Since our sources of recruitment
of participants were Network or newspaper, there may be exited Recruitment bias, like the participants we recruited in this study are more than 60 years old.

5. Conclusion

Over-diagnosis and overtreatment of TN are highly discussed by many research teams in many countries. Many follow-up studies have been conducted on this topic. To explore awareness of over-diagnosis and overtreatment, our study shows participants’ preferences on each criterion, i.e., they thought unmet needs and disease severity are important criteria, but they gave less scores for both criteria regarding TN. Especially physicians’ group gave less score for disease severity than other participants. For citizens’ group who had no experiences and knowledges of this disease, their thoughts of severity are much more severe than the two other groups. Physician indicated keeping the intervention safe and effective were much more important, patients indicated the quality of life after receiving intervention were much more important. However, we cannot roughly conclude from our study, there is overdiagnosis and overtreatment of TN. This study provides a new perspective to explore preferences and insights from different groups. Additionally, through comparison among physicians, patients and citizens, differences have been highlighted in the study, which can make further better communication between physicians and patients. This study provided a suppurative decision making for physicians and policy makers when they conduct researches on TN. We hope the result of this study can contribute to improve diagnosis and treatment of this disease, in addition to ensure sustainable and equitable healthcare resources distribution.

6. List Of Abreviations

TN: Thyroid Nodules

MCDA: Multi Criteria Decision Analysis

EVIDEM: EVidence based Decision-Making

PRO PRICARE: Preventing Overdiagnosis in Primary Care

BMBF: the Germany Federal Ministry for Education and Research

FAU: Friedrich-Alexander University Erlangen-Nürnberg

ACE: Adverse Cascade Effects

QoL: Quality of Life

SD: Standard Deviations

VCx: Value Contribution
7. Declarations

7.1 Ethics approval and consent to participate

Not applicable

7.2 Consent for publication

Not applicable

7.3 Availability of data and materials

The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

7.4 Competing interests

The authors declare that they have no competing interests.

7.5 Funding

The project is part of “PRO PRICARE (Preventing Overdiagnosis in Primary Care)” at the Friedrich Alexander-Universität Erlangen-Nurnberg funded by the Germany Federal Ministry for Education and Research (BMBF). The funding body has no influence in the design of the study, and collection, analysis, and interpretation of data and in writing the manuscript.

7.6 Author contributions

PLK and TK conceived of the study idea. PLK and SZ designed the study. LK designed the online-survey and collected the data. SZ analyzed the data and finalized the manuscript. All authors approved the final manuscript.

7.7 Acknowledgements

Not applicable

8. References

1. American Thyroid Association. *Thyroid Nodules*. 2019; Available from: https://www.thyroid.org/thyroid-nodules/.

2. Hoang, J.K., X.V. Nguyen, and L. Davies, *Overdiagnosis of thyroid cancer: answers to five key questions*. Acad Radiol, 2015. 22(8): p. 1024-9.
3. Verburg, F.A., *Is thyroid surgery performed too often in Germany?* Nuklearmedizin, 2015. **54**(03): p. 101-105.

4. Wienhold, R., et al., *The management of thyroid nodules: a retrospective analysis of health insurance data.* Dtsch Arztebl Int, 2013. **110**(49): p. 827-34.

5. Ahn, H.S., H.J. Kim, and H.G. Welch, *Korea's thyroid-cancer “epidemic”–screening and overdiagnosis.* N Engl J Med, 2014. **371**(19): p. 1765-7.

6. Brito, J.P., J.C. Morris, and V.M. Montori, *Thyroid cancer: zealous imaging has increased detection and treatment of low risk tumours.* Bmj, 2013. **347**: p. f4706.

7. Institute of General Practice University Hospital Erlangen. *Adverse Cascade Effects (ACE) – Undesirable cascade effects.* Trigger and progress of clinical treatment pathways in patients with thyroid anomalies. 2017; Available from: https://www.allgemeinmedizin.uk-erlangen.de/en/research/research-projects/adverse-cascade-effects-ace/.

8. Institute of General Practice University Hospital Erlangen. *Preventing Overdiagnosis in Primary Care (PRO PRICARE).* 2017; Available from: https://www.allgemeinmedizin.uk-erlangen.de/en/forschung/pro-pricare/.

9. EVIDEM. *The Framework_EVIDEM 10th edition.* 2019; Available from: https://www.evidem.org/evidem-framework/.

10. Vaccarella, S., et al., *Worldwide Thyroid-Cancer Epidemic? The Increasing Impact of Overdiagnosis.* N Engl J Med, 2016. **375**(7): p. 614-7.

11. Jegerlehner, S., et al., *Overdiagnosis and overtreatment of thyroid cancer: A population-based temporal trend study.* PLoS One, 2017. **12**(6): p. e0179387.

12. UNIPARK. *Umfragesoftware.* 2020; Available from: https://www.unipark.com/umfragesoftware/.

13. McDow, A.D., et al., *Impact of Rurality on National Trends in Thyroid Cancer Incidence and Long-Term Survival.* J Rural Health, 2019.

14. Roman, B.R., L.G. Morris, and L. Davies, *The thyroid cancer epidemic, 2017 perspective.* Curr Opin Endocrinol Diabetes Obes, 2017. **24**(5): p. 332-336.

15. Cho, H.N., et al., *Determinants of undergoing thyroid cancer screening in Korean women: a cross-sectional analysis from the K-Stori 2016.* BMJ Open, 2019. **9**(4): p. e026366.

16. Goldenberg, D., *We cannot ignore the real component of the rise in thyroid cancer incidence.* Cancer, 2019.

17. Wagner, M., et al., *Can the EVIDEM Framework Tackle Issues Raised by Evaluating Treatments for Rare Diseases: Analysis of Issues and Policies, and Context-Specific Adaptation.* Pharmacoeconomics, 2016. **34**(3): p. 285-301.

18. Jimenez, A., et al., *Determining the value contribution of selexipag for the treatment of pulmonary arterial hypertension (PAH) in Spain using reflective multi-criteria decision analysis (MCDA).* Orphanet J Rare Dis, 2018. **13**(1): p. 220.
19. Wagner, M., et al., Applying Reflective Multicriteria Decision Analysis (MCDA) to Patient-Clinician Shared Decision-Making on the Management of Gastroenteropancreatic Neuroendocrine Tumors (GEP-NET) in the Spanish Context. Adv Ther, 2018. 35(8): p. 1215-1231.

20. Wagner, M., et al., Development of a Framework Based on Reflective MCDA to Support Patient-Clinician Shared Decision-Making: The Case of the Management of Gastroenteropancreatic Neuroendocrine Tumors (GEP-NET) in the United States. Adv Ther, 2018. 35(1): p. 81-99.

21. Wahlster, P., et al., Exploring the perspectives and preferences for HTA across German healthcare stakeholders using a multi-criteria assessment of a pulmonary heart sensor as a case study. Health Res Policy Syst, 2015. 13: p. 24.

22. Kieslich, K., Social values and health priority setting in Germany. J Health Organ Manag, 2012. 26(3): p. 374-83.

23. Pendzialek, J.B., D. Simic, and S. Stock, Measuring customer preferences in the German statutory health insurance. Eur J Health Econ, 2017. 18(7): p. 831-845.

24. American Cancer Society. Can Thyroid Cancer Be Prevented? 2019; Available from: https://www.cancer.org/cancer/thyroid-cancer/causes-risks-prevention/prevention.html.

25. American Thyroid Association. THYROID NODULES Selenium may prevent goiter and thyroid nodules 2011; Available from: https://www.thyroid.org/patient-thyroid-information/ct-for-patients/vol-4-issue-7/vol-4-issue-7-p-6-7/.

26. Klaus Kraywinkel, Epidemiologie ausgewählter endokriner Tumoren in Deutschland. 2018.

27. Shi, L.L., et al., Changes in thyroid cancer incidence, post-2009 American Thyroid Association guidelines. Laryngoscope, 2017. 127(10): p. 2437-2441.