Consolidated Stakeholder Feedback
Levothyroxine for patients diagnosed with subclinical hypothyroidism

Stakeholders **1. IBSA Institut Biochimique SA** **2. Curafutura** **3. Santésuisse** **4. Interpharma**	

1. Kommentar zur Forschungsfrage		**Antwort der Autoren/BAG**
1	The following comment does not concern the above question (1.). We inserted the comment below in this field as a specific field for this particular comment is missing in this form. a. 7.2.2.5, line 507–516, Budget impact analysis: Referring in particular to the cardiovascular risk and cardiovascular mortality (line 82–89), how do the authors intend to account for the costs of hospitalizations and deaths in their budget impact analysis?	Costs of hospitalization and death will be accounted in the cost-effectiveness model. Budget impact model will be based/depend on of the cost-effectiveness model.
2	Die Forschungsfrage wird als sinnvoll erachtet, da die Hypothyreose (Schilddrüsenunterfunktion) eine der häufigsten endokrinen Erkrankungen ist und die derzeitigen Erkenntnisse darauf hindeuten, dass ein beträchtlicher Anteil der europäischen Bevölkerung eine subklinische Hypothyreose (SCH) hat, die aufgrund der fehlenden bzw. nicht offensichtlichen, milden Symptome undiagnostiziert bleibt. Trotz der hohen Prävalenz der SCH (3-18 % der erwachsenen Bevölkerung, populationsabhängig) sind die Erkenntnisse über das Screening sowie das Nutzen/Risikoverhältnis einer Behandlung mit Levothyroxin (T4) umstritten. Da die Diagnose einer SCH allein auf biochemischen Labortests basiert und Patienten mit einer SCH meist beschwerdefrei sind und Langzeitrisiken kontrovers diskutiert werden ("Therapie reiner Laborwerte"), muss geklärt werden, in welchen Situationen der Einsatz von Levothyroxin gerechtfertigt ist.	Acknowledged *(See also answer nr. 4)*.

3	Overall, the research question is correctly formulated. With the four pre-defined populations, the open questions are well addressed. We can support the definitive age and TSH-threshold after the preliminary search. Today, different thresholds are used in clinical guidelines. It is therefore necessary to consider them in a full HTA. In particular, the guidelines used and referenced in Switzerland should be considered.	Acknowledged *(See also answer nr. 4)*.
4	Der Fokus der Forschungsfrage ist sehr breit gefasst und es wird kein wirkliches Ziel formuliert, was mit dem HTA erreicht werden soll: - Es sollen vier Populationsgruppen, die das gesamte Spektrum der SCH abdecken, untersucht werden im Gegensatz zur ursprünglich beschriebenen zu evaluierenden Technologie in der Plausibilisierung, bei der auf Patienten > 65 Jahren mit SCH fokussiert wurde. - Es fehlt unsere Ansicht nach eine Begründung warum jede der Subgruppen im Rahmen des HTA analysiert werden soll. - Es werden keine definitiven Alters- und TSH-Grenzwerte festgelegt. Diese werden einen erheblichen Einfluss auf das HTA-Assessment haben. Eine Festlegung der Grenzwerte ohne Einbezug von Schweizer Experten und den medizinischen Verbänden sehen wir als kritisch. - Ist es ethisch vertretbar Vergütungsentscheide auf einem klar definierten Patientenalter in Betracht zu ziehen? siehe auch Anhang Begleitbrief	The original application for this topic indeed focused mainly on the elderly population. However, the treatment of SCH with Levothyroxine in patients below the age of 65 is also controversially discussed in the literature. (See for example Cochrane review: Thyroid hormone replacement for subclinical hypothyroidism (nih.gov) or NICE HTA: NG145 Evidence review G (nice.org.uk). Also in medical practice, it remains a controversial topic as seen by the clinical case studies discussions: To Treat or Not to Treat Subclinical Hypothyroidism, What Is the Evidence? (nih.gov) and Subclinical hypothyroidism: to treat or not to treat? in: European Journal of Endocrinology Volume 183 Issue 6 (2020) (bioscientifica.com). Consulted experts for this project have confirmed this controversy and advised the HTA team to investigate different age groups within the general SCH patient population. Given the apparent confusion the current formulation of the research question has caused, the research question has been reformulated. The suggested stratification thresholds by age and TSH levels have been removed. Gender, age and TSH data will be extracted from the included studies. After the

		data-extraction, the possibilities for stratification of the outcomes on gender, age and/or TSH level will be investigated. The ethical aspects associated with coverage limitations based by age will be discussed in the HTA report.
5	In Ergänzung zum Formular: Die Research question hinterfragt die in Behandlungsleitlinien empfohlenen Grenzwerte, die von Experten aufgrund der zur Verfügungstehenden klinischen Evidenz erarbeitet werden, erneut. In der TRUST-Studie (siehe Plausibilisierung: «Gründe für ein HTA: Die Ergebnisse der im Frühjahr 2017 publizierten TRUST-Studie mit Schweizer Beteiligung zeigen auf, dass bei der subklinischen Hypothyreose keine Intervention eine Option sein kann.») wurden nur Patienten >65 Jahren mit SCH untersucht. Gemäss HTA-Protokoll sollen nun vier Subgruppen analysiert werden, die das gesamte Spektrum der SCH abdecken, auch Subgruppen, in denen eine Therapie mit Levothyroxin unbestritten ist.	The normal value range from TSH is not being questioned in this HTA report. Amendments are made to avoid this suggestion. The original application for this topic indeed focused mainly on the elderly population. However, the treatment of SCH with Levothyroxine in patients below the age of 65 is also controversially discussed in the literature. (See for example Cochrane review: Thyroid hormone replacement for subclinical hypothyroidism (nih.gov) or NICE HTA: NG145 Evidence review G (nice.org.uk). Also in medical practice, it remains a controversial topic as seen by the clinical case studies discussions: To Treat or Not to Treat Subclinical Hypothyroidism, What Is the Evidence? (nih.gov) and Subclinical hypothyroidism: to treat or not to treat? in: European Journal of Endocrinology Volume 183 Issue 6 (2020) (bioscientifica.com). Consulted experts for this project have confirmed this controversy and advised the HTA team to investigate different age groups within the general SCH patient population. Given the apparent confusion the current formulation of the research question has caused, the research question has been reformulated. The suggested stratification thresholds by age and TSH levels have been removed. Gender, age and TSH data will be extracted from the included studies. After the data-

		extraction, the possibilities for stratification of the outcomes on gender, age and/or TSH level will be investigated.

2. Kommentar zum PICO		
1	a. 5, Line 130, Table 1: PICO, Population: The authors intend to stratify the patients according to age and TSH level and to define the definitive age and TSH level thresholds after a preliminary data extraction of identified RCTs. Such an approach will lead to the exclusion of several RCTs. We, therefore, suggest that the authors carry out sensitivity analyses to capture the information lost by such an exclusion. Assuming for example that the authors decided to use an age of 65 years and a TSH level of 8 mU/I as thresholds, they should pool in the sensitivity analyses the at first excluded RCTs with the thresholds: (i) 70 years/TSH level 8 mU/l; (ii) 65 years/TSH level of 10 mU/l; (iii) 70 years/TSH level of 10 mU/l. b. 5, Line 130, Table 1: PICO, Outcome, Efficacy/effectiveness outcomes: We also suggest that the authors identify all available validated scores, make a clear decision regarding the scores they intend to use and rank them hierarchically according to clinical relevance.	a. During the full-text selection all studies in patients diagnosed with SCH in which levothyroxine is compared with placebo will be included, irrespective of age and TSH levels/thresholds. After the data-extraction, the possibilities for stratification of the outcomes on gender, age and/or TSH level will be investigated. b. We will extract all available data on these scores.
2	Es sollte geprüft werden, ob neben den Altersklassen und TSH-Werten eine Unterscheidung des Geschlechts in die Forschungsfrage mitaufgenommen werden sollte (siehe dazu Kommentar unter Punkt 4). Frauen weisen eine deutliche höhere Prävalenz für eine SCH auf.	Many factors influence the TSH levels. Clinical guidelines for the general population only use age and TSH thresholds for the indication of levothyroxine treatment for SCH. After the data-extraction, the possibilities for stratification of the outcomes on gender, age and/or TSH levels will be investigated. Studies only including women with specific female sex hormonal states, e.g. pregnant women, non-pregnant women on fertility-related treatment, or menopausal women are out of scope for the current HTA.
3	It is intended to consider studies where products with Levothyroxin are compared to a Placebo or no treatement. We recomend also considering studies where different products (brands) and / or different doses with Levothyroxin are compared. Since there are differences between the products this could provide important additional information. It is not clear why studies comparing Levothyroxin with other drug treatments should be excluded a priori.	Comparison between medication doses or drug-drug comparisons are out of scope.
4	P: Die Wahl der vier Patientengruppen sowie die vorgesehenen Grenzwerte sind anderweitig zu erläutern. Es ist auf die Subgruppe zu fokussieren, wo der Einsatz in Frage gestellt wird (siehe Plausibilisierung: Patienten >65 Jahren). Eine Festlegung	P:. See answer nr.4 from section 1. O: Withdrawal of treatment due to lack of efficacy of

	der Grenzwerte erst nach Beginn der Literaturrecherche deutet auf ein unklares Ziel der gesamten Analyse hin und stellt die Qualität der systematischen Recherche in Frage. Es ist zu begründen, weshalb nicht nur Alter, sondern auch TSH-Grenzwerte zur Wahl der Patientengruppen erst im Verlauf der Suche definiert werden sollen. O: Efficacy O: der Unterpunkt «Withdrawal of treatment due to lack of efficacy of levothyroxine" ist kein Effektivitätsoutcome. Safety O: eine genauere Definition der "serious and clinically important AE" sollte möglich sein Economic O: es fehlen gesetzliche Vorgaben bzgl. ICER und Budget Impact	levothyroxine is related to the efficacy outcomes. This outcome should not be confused with withdrawal of treatment due to safety issues. Inherent to the nature of any subclinical disorder the inclusion of subjective outcomes such as *withdrawal of treatment due to lack of efficacy of the treatment* in addition to the more clinical outcomes is considered very relevant. Safety O: all data reported on serious and clinically relevant AEs will be extracted from the included studies. Economic O: The methodology applied to develop an economic model will be described in detail in the HTA report. The lack of a predefined national cost-effectiveness or willingness-to-pay thresholds does not preclude calculating different incremental cost-effectiveness ratios, by assuming reasonable thresholds. All study findings will be interpreted with the necessary caution and will be analysed in various sensitivity analyses.

3. Kommentar zu Datenbanken und Suchstrategie		
1	a. 7.1.1.1, Line 181–183: According to Muka et al. (Eur J Epidemiol 2020;35 (1):49–60) a literature search (i.e., a search strategy) should include at least four online databases to ensure adequate coverage. Several databases can be considered, including PubMed, EMBASE, Cochrane Library, ClinicalTrials.gov, Web of Science, CENTRAL, Emcare, etc. – The expected overlap does not warrant the proposed limitation to just two databases.	The choice for the literature databases was discussed and agreed upon by the FOPH project team. See also answer to comment 4.
2	Die Beschreibung der Suchstrategie und des Auswahlverfahrens der Studien ist sinnvoll und methodisch korrekt.	Thank you.

3	Santésuisse would like to suggest that, in addition to searching the PubMed and Embase.com databases, the ClinicalTrials.gov, Embase, Google Scholar and HTAiVortal databases should be taken into account. In addition, it is recomended to look for ongoing research on this topic. It would give an idea whether new results can be expected in the near future. We also propose to search in the third Swiss language, Italian.	- The choice for the literature databases was discussed and agreed upon by the FOPH project team. It was chosen not to expand the search in additional databases, because RCTs on medication are sufficiently covered with the two databases PubMed (MEDLINE) and Embase.com. Clinicaltrials.gov is a database of privately and publicly funded clinical studies conducted around the world and does not contain peer-reviewed literature and therefore will not be included as database for the clinical systematic review. Relevant ongoing trials (including those on ClinicalTrial.gov) will be addressed in Chapter 11 of the HTA report. Google Scholar is not a reliable source for systematic reviews, because amongst others the search results are not reproducible. HTAiVortal is not a comprehensive database. HTAi is an international society and a conference website which can be considered for grey literature search. - Peer-reviewed articles reporting preliminary results of ongoing studies will be included if applicable. - The choice for languages was discussed and agreed upon by the FOPH project team.
4	L167: Schritt I wird in Z.179 bereits wieder gestrichen ohne transparente Darstellung warum. L183: Es ist zu erklären, warum die Cochrane Library nicht berücksichtigt wird L189: Die Auswahl der Suchsprachen ist zu begründen. L205-207: Auswahl der Studien ohne klar definierte Kriterien L208-210: Vorgehensweise data extraction phase nicht klar definiert L221: 'Relevante Kriterien' sind im Voraus zu definiert unter Einbezug der Stakeholder. L224: "study quality": Please define "sufficient methodological quality", how will it be assessed?; "study population": Would it be beneficial to specifically mention ft4 levels? If not, please explain somewhere in the protocol why ft4 levels are not part of the PICO or the in-/exclusion criteria.;"study intervention": follow up duration should be predefined e.g. lifetime;"short treatment or follow-up duration (to be determined in further detail during the project)" Please describe on what this decision will be based. siehe auch Anhang Begleitbrief	L167: This is explained in line 176-180. L183: It was chosen not to search in Cochrane Library as additional database, because RCTs sufficiently covered with the two databases Medline (i.e. searched using PubMed) and Embase. Besides, the RCTs reported in Cochrane library are retrieved from Medline and Embase. L189: The choice for languages was discussed and agreed upon by the FOPH project team. L205-207: These criteria are enclosed in Chapter 7.1.1.3; as referred to at the end of line 207.

		L208-210: This section describes the screening during data extraction phase, the data extraction is outlined in section 7.1.2.1. L221: When gaining more insight in the titles/abstracts of the studies, it might be that an additional criterion comes forward which needs to be explicitly mentioned in the inclusion/exclusion table. This will only be added after careful discussion with the FOPH and will be reported in the HTA report. L224: - Study quality: If studies are excluded based on major methodological flaws, the specific reason for studies excluded will be provided in the HTA report. We want to highlight that studies are rarely excluded because of this reason. If there is a methodological issue, most studies will be included and the methodological flaw will be mentioned in the risk of bias assessment. - Study population: The study population is defined as 'Patients diagnosed with SCH'. For the diagnosis of SCH, TSH and fT4 levels are necessary (as explained in the medical background). The fT4 is therefore already incorporated within the PICO and inclusion/exclusion criteria. - Study intervention: we will not limit the selection at beforehand on follow-up duration. When choosing lifetime as follow-up duration, all RCTs will be excluded based on this selection criterion. - Short treatment or follow-up duration (to be determined in further detail during the project): if studies with an outlier short treatment duration in comparison with the remaining studies are found, it can be decided to define an additional exclusion criterion for treatment duration (e.g., an outlier treatment duration of 1 week versus treatment durations of 6 to 12 months).

4. Kommentar zu Datenextraktion, Analyse und Synthese		
1	a. 7.1.1.3, Line 224, Table 2, Country of study, and line 230/231: We do not agree with the exclusion of 'all other countries'. Iodine nutrition policies may have an impact on the incidence of SCH, but not on the research question (line 43–45). b. 7.1.1.4, 7.1.2.1 and 7.1.2.2, line 233–306: We carried out a search in PubMed and came to the conclusion that the literature available to answer the research question is not very ample. The outlined step-by-step approach, therefore, is not justified. The selection process ('include in' or 'exclude from' the 'full text assessment'), the full text-assessment itself as well as the data extraction (7.1.2.1) and the critical appraisal (7.1.2.2) thus always ought to be carried out independently by two reviewers and resolved by a third one in case of discrepancies as outlined in 7.1.2.4 (double-check of all steps).	a. We agree with the reviewer and changed the inclusion criterion to all countries. b. - The stepwise approach is based on the level of evidence. The focus of HTAs conducted by the FOPH is to search for the highest available scientific and most homogeneous evidence provided by RCTs. In case no RCT or only one RCT is found, an additional systematic literature search will be conducted for comparative non-randomised studies. - The implemented quality control measures are agreed upon by the FOPH project team.
2	Weshalb werden Daten ausgeschlossen, welche explizit weibliche Populationen untersuchen, insbesondere menopausale Frauen? --> Exclusion criteria for RCT and comparative non-randomised studies with study populations: "studies including women with specific sex hormonal states, eg. menopausal women" Leistungszahlen der Versicherer zeigen hierzu, dass Levothyroxin-Bezüger in 80.8% der Fälle weiblich sind. Insbesondere sind 72.9% der weiblichen Levothyroxin-Bezüger älter als 50 Jahre. Deshalb sollten keine Daten oder Studien aufgrund des Geschlechts und des Alters der Studienpopulation ausgeschlossen werden.	Studies with mixed study populations including women are of interest. Studies focusing on a specific population (e.g., only including women with specific female sex hormonal states affecting the TSH levels) are out of scope.
3	We strongly welcome the use of the GRADE instrument to assess the quality of evidence. The scoping addresses the relevant questions. santésuisse supports the implementation of a full HTA. A new cost-benefit analysis based on the figures and framework conditions in Switzerland is expedient.	Acknowledged. No amendment needed.

4	L271: This spreadsheet should be pretested. L362: will recommended filters for health economic studies be used? Please clarify. L370: Was ist die Definition von 'relevant data'? L244/L397: die Expertise der Forscher ist offen zu legen L365-379: By how many reviewers? All in parallel? Please clarify. L397: What about title/abstract? Will they be reviewed only by one researcher? L447-L520: Very generic. Could be part of any protocol independent of the HTA topic. L436: zu einem de novo gesundheitsökonomischen Modell sollte erneut ein Stakeholdereinbezug erfolgen; es gibt keine behördlichen Vorgaben in der Schweiz für die Erstellung eines solchen Modells L487: We would recommend using the Swiss Consumer Price Index for the healthcare sector to adjust for changes in prices over time. L518-520: How are these criteria going to be analyzed based on the included studies?	L271: A first set of studies will be extracted in the Excel spreadsheet and finetuned in detail by two researchers, then the remaining of the included studies will be extracted according to this format. L362: The recommended filters will be used. L370: Data considered appropriate based on inclusion and exclusion criteria. L244/L397: At this point in the HTA process, the qualifications of the HTA team will not be shared with the stakeholder. However, the FOPH guarantees that the qualifications of the HTA team fulfil all requirements for pursuing this HTA report. L365-379: As indicated in the quality control section, all steps will be performed and checked by two independent reviewers. In case of discrepancy, a third researcher will be consulted. L397: As indicated in the quality control section, all steps will be performed and checked by two independent reviewers. In case of discrepancy, a third researcher will be consulted. L447-L520: This is similar to previous sections of the HTA protocol. HTA report will contain all of the specific details/explanations for this topic. L436: The methodology applied to develop an economic model will be described in detail in the HTA report. The lack of a predefined

		national cost-effectiveness or willingness-to-pay thresholds does not preclude calculating different incremental cost-effectiveness ratios, by assuming reasonable thresholds. All study findings will be interpreted with the necessary caution and will be analysed in various sensitivity analyses. L487: We will implement your feedback in the HTA report. L518-520: We will perform grey literature searches on these HTA domains. In the HTA report, we will identify and report the main legal, social, ethical, and organisational issues of the studies included in the efficacy, effectiveness, and safety and cost-effectiveness systematic literature searches.

5. Kommentar zu Policy Question		
1	L31-42: Please provide a specific policy question. This chapter rather describes the background than a policy question. A specific policy question would be of real benefit to this HTA as age of the patients and serum TSH levels may play an important role but might be rather challenging to include in policies. L37: The HTA should investigate the meaning of the referenced clinical guidelines in Switzerland. Please consider Swiss specific publications like the one by Fischli (Hypothyreose. Schweiz Med Forum, 2013. 13(21): p. 401-407) or the medix guideline on “Schilddrüsenerkrankungen”. Es fehlen wesentliche Ausführungen aus der Plausibilisierung, warum dieses HTA durchgeführt werden soll, obwohl das Einsparpotential gemäss der Schlussbewertung des BAG als gering eingeschätzt wird (u.a. «Die Ergebnisse der im Frühjahr 2017 publizierten TRUST-Studie zeigen auf, dass bei der subklinischen Hypothyreose keine	L31-42: The policy question has been reworded. L37: The guidelines will be discussed in the HTA report. The references will also be shared with the HTA thyroid hormone team.

	Intervention eine Option sein kann»; «Seit 2010 hat der Umsatz der Levothyroxin-Präparate um rund 40% zugenommen.»). Ebenfalls müsste auf den Zusammenhang des parallel laufenden HTA- Themas in Bezug auf den Einsatz von Schilddrüsenhormontests und der beobachteten Zunahme der Thyroid Funktionstests in den letzten Jahren und diesem Verfahren zum Einsatz von Levothyroxin eingegangen werden.	