Consolidated Stakeholder Feedback

Thyroid function tests for the diagnosis of suspected primary or secondary thyroid dysfunction

Stakeholders
1. Curafutura
2. Santésuisse
3. SVDI
4. FAMH
5. Société médicale du Valais (SMVS)

1. Kommentar zur Forschungsfrage		**Antwort der Autoren/BAG**
Curafutura	Für die Diagnose von Schilddrüsenerkrankungen empfehlen die Leitlinien ein 2-stufiges Testverfahren: TSH-Messung, wenn TSH ausserhalb des Referenzbereichs liegt, folgt nachgelagert eine Messung von T4/T3 aus derselben Probe. Es sei beschrieben, dass TSH & T4/T3 oft zusammen in einem 1-stufigen Testansatz gemessen werden. Das Argument für den 1-stufigen Ansatz ist die Vermeidung von Fehldiagnosen bei Patienten mit normalen TSH-Werten aber T4/T3 Werten ausserhalb des Referenzbereichs. Die Kosten für die Bestimmung von T4 und T3 sind überschaubar und eine erneute Auftragstaxe sogar teurer. Die Frage sollte aus unserer Sicht schon früher ansetzen: Wann macht eine Laborabklärung auf die Parameter überhaupt Sinn? Liegen richtungsweisende Symptome/Befunde vor, welche abgeklärt werden müssen? Liegt ein akzeptable Wahrscheinlichkeit für eine therapeutische Konsequenz vor? Weil: Ca. 89% der TSH-getesteten Patienten bleiben ohne medikamentöse Therapiekonsequenz (keine Medikation mit ATC H03)	These questions are indeed relevant within the context of endocrine disorder diagnostics. However, some of these questions are clearly guideline issues rather than HTA issues and are beyond the scope of this HTA.

Santésuisse	The protocol explains clearly the research questions.	Thank you.
SVDI	Grundsätzlich erscheint uns die klinisch-wissenschaftliche Forschungsfrage als akkurat und wichtig. Ganz generell ist T3, T4 nicht mehr Standard und es wird in verschiedenen Ländern eine sinnvolle Stufendiagnostik empfohlen (z.B. TSH und ft4, danach wird weiter differenziert) Unsere Anmerkungen zur Forschungsfrage: 1. Testanforderung: Ist der Einschritt-Ansatz tatsächlich immer einer Anforderung aller Tests oder ist dies nicht mehr zeitgemäss und wird entsprechend angepasst (z.B. Reflex nach TSH hoch/niedrig)? 2. Reflex Testung: Spielt das Thema automatisierter Reflextests eine Rolle oder erfordert der erste Schritt immer erst eine ärztliche Bewertung und Anforderung der Folgetests? Folgetests haben durchaus weitere Kosten zu Folge (Auftragspauschalen im Labor, ärztliche Leistungen etc.) 3. Wie wird damit umgegangen, dass TSH sehr variabel ist und daher eine TSH Bestimmung zwei mal durchgeführt werden soll, zu den gleichen Bedingungen aber an unterschiedlichen Tagen?	We acknowledge that T4/T3 testing is no longer the standard and have changed the text in the protocol from T4/T3 testing to (f)T4/(f)T3 testing. Please be aware that the focus of this HTA is on the one-step test approach compared to the two-step test approach. The possible combinations of thyroid hormones tests to identify thyroid dysfunction in suspected adults is relevant background information but not the focus of this HTA. Additional information on the appearance of total T4/T3, free T4/T3 and reverse T3 in blood and general information on thyroid hormone testing is added to the medical background. 1. In this HTA we define the one-step test approach as measurement of TSH and (f)T4/(f)T3 simultaneously and not as reflex ((f)T4/(f)T3 testing performed automatically based on abnormal TSH result. To identify patients with normal TSH and abnormal (f)T3/(f)T4 (i.e. one of the primary clinical outcomes), we need studies that applied the one-step test approach. Studies with only reflex testing/a two-step test approach will not identify these patients and will therefore not be included in the clinical systematic review. 2. This is an organisational issue, for which expert clinicians will be consulted in the HTA phase. 3. We will extract this type of data as reported in the included studies. Also, we will address this in the chapter of organisational issues of the HTA report.
FAMH	Il s'agit d'une question pertinente sur le fond. Cependant, 2 points clarifications sont impératifs: 1) Le chiffres exacts de la répartition des bilans thyroïdiens séquentiels versus complets en Suisse manquent. En leur absence la pertinence et l'utilité de ce projet HTA semble proche de zéro. En effet, à l'heure les test immunologiques sont utilisés de manière vastement majoritaire et où les recommandations internationales par ailleurs citées, la grande majorité dosages de première intention se font déjà par tests réflexes et répondent aux besoins (le risque de rater une dysthyroïdie centrale via la stratégie séquentielle étant faible et rattrapable). 2) Ce dossier compare des indications/modalités de prescription, point clé pour la FAMH puisque les laboratoires ne maîtrisent pas la	1) An increase in the number of thyroid function tests is observed, and often TSH and (f)T4/(f)T3 are measured together in a one-step test approach. From a clinical point of view, the only differences between a one-step test approach and a two-step test approach is missing patients with a normal TSH and abnormal (f)T4/(f)T3. With the clinical systematic review, we want to gain insight on the number of diagnosis missed in adults with normal TSH and abnormal (f)T4/(f)T3 levels among those suspected of thyroid disorder, as reported in peer-reviewed literature. From an economic perspective, we'll quantify the costs and budget impact of these tests. 2) HTA is a method of evidence synthesis that considers evidence regarding clinical effectiveness, costs, budget impact

	prescription des analyses. Hors, il n'existe à l'heure actuelle aucune démonstration de l'efficacité d'un processus HTA/impact sur le remboursement sur les pratiques de prescription.	and, when broadly applied, includes social, ethical, organizational and legal aspects of the use of health technologies. A major use of HTA is in informing reimbursement and coverage decisions by insurers and national health systems. This process is different from clinical guideline prescription and/or recommendation.
SMVS	Nous retenons de cette étude les objectifs suivants : combien de diagnostics seraient manqués lors d'investigation de troubles thyroïdiens dans le cas d'investigations en une étape (dosage simultané TSH et T3 / T4) versus en deux étapes (dosage T3 / T4 dans un second temps si TSH hors norme), quels seraient les diagnostics manqués, enfin quel est l'impact économique de ces méthodologies. Si la question véritable est l'économicité de tels traitements, toute conclusion, plus encore si des directives devaient être imposées, va se heurter aux recommandations cliniques, au détriment d'une cohérence thérapeutique. Les directives relatives à la prise en charge médicale doivent rester du ressort des sociétés de discipline et des "opinion leader", afin de garantir le libre choix thérapeutique du médecin.	HTA is a method of evidence synthesis that considers evidence regarding clinical effectiveness, costs, budget impact and, when broadly applied, includes social, ethical, organizational and legal aspects of the use of health technologies. A major use of HTA is in informing reimbursement and coverage decisions by insurers and national health systems. This process is different from clinical guideline prescription and/or recommendation.

2. Kommentar zum PICO		
Curafutura	Kommentar zu PICO 2: Outcome: Es müsste eruiert werden, wie hoch der Einsparungsbetrag sein müsste, damit ein Qualitätsverlust in der Diagnostik durch das einstufige Verfahren (und um welches Delta es sich handeln darf) vertretbar wäre (analog zum ICER - Modell, Quadrant "weniger gut, aber billiger").	Thank you for your feedback. We will do both costs and budget impact analyses and quantify incremental savings.
Santésuisse	Both PICO (one for the clinical systematic review and one for the cost and budget impact analysis) are in line with the medical background as well as the technology description. However, there seems to be a discrepancy between the PICO for the clinical systematic review and the inclusion / exclusion criteria concerning the defined population. If this difference is deliberate, it should be precisely explained and justified. This discrepancy does not exist in the part of the cost and budget analysis.	The title and both PICO's are adapted and aligned with the inclusion/exclusion table

SVDI	- Im Titel steht "Thyroid function tests for the diagnosis of suspected primary hypo- or hyperthyroidism" jedoch werden bei Tabelle 1 alle Formen aufgezeigt und auch im PICO werden alle Kombinationen an TSH, T3, T4 Level angegeben. Weshalb werden alle Kombinationen angegeben (gesucht), wenn sich die Fragestellung nur auf primäre Über- oder Unterfunktion der Schilddrüse bezieht? - Es wird nicht auf die Frage eingegangen, wann es sinnvoll ist, das totale T3 bzw. T4 oder das freie T3,T4 oder Kombinationen davon zu messen. Je nach Situation ist das eine oder andere mehr angebracht (siehe auch Kommentare zu Fragen 1 und 4). Auch wird das reverse T3 nicht erwähnt.	- Thank you for your comment, the research question is changed to adults with suspected primary or secondary thyroid dysfunction. We are looking for all possible diagnoses that would have been detected with the 1-step testing approach (or missed diagnoses with the 2-step testing approach) in this population. - The focus of this HTA is on the one-step test approach compared to the two-step test approach. The possible combinations of thyroid hormones tests to identify thyroid dysfunction in suspected adults is relevant background information, but not the focus of this HTA. Additional information on the appearance of total T4/T3, free T4/T3 and reverse T3 in blood and general information on thyroid hormone testing is added to the medical background. This kind of data (i.e. which thyroid hormone tests were performed to diagnose thyroid dysfunction) will be extracted from the included studies.
FAMH	1.Description, lignes 134/135: dire que le test séquentiel nécessite deux prélèvements est faux, tout est fait sur le même échantillon (mais de manière séquentielle) en cas de TSH anormale. Ce point doit être précisé/corrigé. 2.Comparateur pour la revue de la litérature: not applicable (!): Ne pas prendre de comparateur qui en l'occurence devrait être les données existantes dans la litérature des coûts/impacts d'une stratégie non séquentielle (les 3 test en même temps) est difficilement compréhensible et ne renforcera pas la crédibilité des conclusions de ce dossier. 3. La population doit être prise en compte dans la stratégie de recherche pour la deuxième question -> Combien de personnes sont dépistées annuellement pour la TSH (et ne sont pas suivies) ? Ou combien ont un soupçon de prim. Hypo-/Hyperthyréose ? Cela ne correspond pas à l'incidence.	1. We modified the sentence as follows: The test is done sequentially on the same sample in the event of abnormal TSH. 2. When adding a comparator to the PICO, many studies with relevant data have to be excluded. There are hardly any studies that directly compare the 1-step test approach with the 2-step test approach for the diagnosis of thyroid disorders, as both tests (TSH and (f)T3/(f)T4) are needed to diagnose a thyroid disorder. With data from the 1-step test approach, cases missed with the 2-step test approach can be identified. 3. Screening is out of the scope of this HTA. However, misuse or overuse of thyroid function tests will be addressed in the organizational issues chapter of the HTA report.
SMVS	Nous ne mettons pas en doute le sérieux de la méthodologie et les critères utilisés du PICO, qui avec ses 9 catégories couvre l'ensemble des profils potentiels, mais ne peut aussi strictement refléter l'hétérogénéité des patients qu'un médecin est appelé à voir dans sa	We added a sentence to the medical background to highlight the heterogeneity of the patients/indications for testing and will address this issue in the discussion of the HTA report. This being said, an HTA provides the available evidence related

	consultation. Les questions auxquelles cette étude veut répondre ne sont pas applicables, tout au moins si les conclusions iraient dans le sens d'une restriction des choix thérapeutiques du médecin. Des recommandations sont souhaitables afin de garantir la qualité de toute prise en charge dans un souci d'efficience, mais ces recommandations doivent être issues d'un véritable objectif de qualité et de rationalité.	to the research question. An HTA provides the assessment and analysis of the data. The HTA does not contain an appraisal judgment of the data.

3. Kommentar zu Datenbanken und Suchstrategie		
Curafutura	Das geplante Vorgehen erscheint methodisch korrekt und sinnvoll.	Thank you.
Santésuisse	Santésuisse suggests 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 the context of the research question, it is important to also include existing guidelines or at least give an overview of them, as well as the current procedures (not only in Switzerland but also in other countries). There is little on this in the present report.	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.
SVDI	Exclusion Criteria: aus unserer Sicht sollten Schwangere ausgeschlossen werden von dieser Analyse.	Agreed. We added “Studies only including women with specific female sex hormonal states, e.g. pregnant women, non-pregnant women on fertility-related treatment, or menopausal women” as exclusion criterion
FAMH	- La plupart des données sur les dysthyroṙidie et leur diagnostique différentiel datent d'avant 1995. Il faut étendre la période à 1990. Dans ces années précédent l'ère de l'EBM, la qualité de données ne répond	- The focus of the HTA is on two different testing approaches (i.e. one-step testing approach and two-step testing approach) and not on dysthyroidism (or any outcome specifically). The

	souvent pas aux critères actuels bien que notre pratique médicale actuelle en découle. Une stratégie de mitigation de ce point doit être proposée. - pourquoi l'italien et non l'espagnol (-> nombre de publications pertinentes) ? - Tab. 4 : Study population exclusion -> Pas d'adultes sous traitement, pas de sec. hypo/hyperthyroïdie - Pour l'analyse de l'impact budgétaire, le nombre de cas suspects par an doit être considéré comme indiqué dans le PICO. Sinon, seuls les coûts relatifs peuvent être évalués.	time period 1995 to 2021 is based on the publication dates found with a preliminary search for key articles relevant for this HTA, and discussed and agreed upon by the FOPH project team. Nonetheless, it was decided to extend the search period from 1995-current to 1990-current to cover a broader field of possibly relevant articles. - Italian is a national language, and all national languages are included. This is based on stakeholder consensus. - No treatment was indirectly included in the inclusion criterion for study population, adults on treatment are not suspected cases but already diagnosed with a thyroid disorder. We added ‘populations monitored for the treatment of thyroid disorders’ as exclusion criterion. - For the budget impact analysis, the number of suspected cases per time (e.g., year) will be considered.
SMVS	Comment justifier que les médecins suisses, et surtout les patients, seraient privés d'une liberté de choix diagnostique qui fait partie des options reconnues ? Comment justifier qu'une étude pose la question du caractère "effectiveness" d'examens de laboratoire à priori peu onéreux, souvent indispensables, au risque de manquer des diagnostics et de devoir reconvoquer le patient, avec des coûts bien évidemment plus élevés. ? Accepter une telle limitation pose la question, avec gravité, du prix que l'on est prêt à accepter pour un traitement efficace et reconnu. Une limitation entrerait en contradiction avec les études cliniques et recommandations internationales et nationales concernant la prise en charge des troubles thytoïdiens.	HTA is a method of evidence synthesis that considers evidence regarding clinical effectiveness, costs, budget impact and, when broadly applied, includes social, ethical, organizational and legal aspects of the use of health technologies. A major use of HTA is in informing reimbursement and coverage decisions by insurers and national health systems. This process is different from clinical guideline prescription and/or recommendation.

4. Kommentar zu Datenextraktion, Analyse und Synthese		
Curafutura	Das geplante Vorgehen erscheint methodisch korrekt und sinnvoll.	Thank you.

Santésuisse	Santésuisse appreciates the assessment of the budget impact. In addition, a cost-benefit analysis of the intervention should be carried out.	The classical HTA approach is not suited for the objective of this HTA report. The central research questions focus mainly on the number of missed adults with suspected primary or secondary thyroid dysfunction when applying the two-step approach and the clinical and economic consequences of the one-step compared to the two-step approach. Therefore, an adapted HTA approach will be used to focus on costs and budget impact analyses.
SVDI	1. Standardisierung: Das Studienkonzept hebt die Problematik einer fehlenden Standardisierung oder Harmonisierung hervor. Wie wird mit dieser Problematik umgegangen, welche Techniken werden angewendet um diesen Bias zu kontrollieren / auszuhebeln? 2. Subgruppen: Gibt es im Rahmen der sehr klar umrissenden Fragestellung einige (Risiko-) Gruppen, die besonders hervorzuheben wären und bei denen ein Einschritt-Ansatz vorteilhafter sein kann als in anderen Gruppen? 3. Wird bei der Analyse auf Bereiche eingegangen wie das Neonatalscreening, Hormonspiegel, oder Verlaufskontrolle nach Therapiemassnahmen? 4. Gibt es eine Differenzierung im Hinblick auf den Algorithmus für stationäre Patienten und ambulante Patienten? 5. Für die Ausführung der Stufen-Diagnostik ist es sicherlich sehr wichtig, dass die zuständige Ärzteschaft ein sehr gutes Verständnis hat von der Thyroid Funktionsdiagnostik, damit auch subklinische Fälle detektiert werden.	1. We will extract the data on this topic as reported in the included studies. The options for clinically relevant data merging/stratification will be discussed with clinical experts, based on the level of detail of data reporting in the studies. 2. We will extract the data on this topic as reported in the included studies. The options for clinically relevant data merging/stratification will be discussed with clinical experts, based on the level of detail of data reporting in the studies. 3. Screening and follow-up after therapy measures are out of scope. The difference between screening studies and diagnostic studies is explained at the end of the medical background. This is indirectly included in the definition of the population (adults with suspected primary or secondary thyroid dysfunction). Based on this comment, we added screening and monitoring as exclusion criteria. 4. We will extract the data on this topic as reported in the included studies. The options for clinically relevant data merging/stratification will be discussed with clinical experts, based on the level of detail of data reporting in the studies. 5. Acknowledged. No amendment needed. We will incorporate this feedback in the organisational issues section of the HTA report.
FAMH	La plupart des données sur les dysthyroridie et leur diagnostique différentiel datent d'avant 1995. Il faut étendre la période à 1990.	The focus of the HTA is on two different testing approaches (i.e. one-step testing approach and two-step testing approach) and not on dysthyroidism (or any outcome specifically). The time period 1995 to 2021 is based on the publication dates found with

		a preliminary search for key articles relevant for this HTA, and discussed and agreed upon by the FOPH project team. Nonetheless, it was decided to extend the search period from 1995-current to 1990-current to cover a broader field of possibly relevant articles.
SMVS	L'étude actuelle pose la question avouée de la pertinence d'effectuer simultanément ou en deux étapes les tests thyroïdiens, mais aussi de l'économicité de tels tests dans la prise en charge des patients avec suspicion de dysthyroïdie. En acceptant ce type d'étude, on pourrait voir le risque de remettre en question de multiples thérapies. Ce genre d'étude est un réel danger de voir un système imposer aux médecins des choix thérapeutiques sur des impératifs financiers, et ses conclusions vont se heurter aux recommandations cliniques, au détriment d'une cohérence clinique et du respect des recommandations. Nous nous permettons ainsi de conclure que ce type de "scoping report" HTA concernant des examens reconnus n'est pas adapté ni adaptable à la clinique et ne doit pas être reconduit pour d'autres examens, car il n'est pas adéquat d'utiliser des données secondaires pour définir l'économicité d'un traitement.	HTA is a method of evidence synthesis that considers evidence regarding clinical effectiveness, costs, budget impact and, when broadly applied, includes social, ethical, organizational and legal aspects of the use of health technologies. A major use of HTA is in informing reimbursement and coverage decisions by insurers and national health systems. This process is different from clinical guideline prescription and/or recommendation.