Developing a Core Outcome Set for clinical trials in Olfactory Disorders: a COMET initiative

C. Philpott, K. Kumaresan, AW. Fjaeldstad, A.Macchi, G. Monti, J Frasnelli, I. Konstantinidis, J. Pinto, J. Mullol, J. Boardman, J. Vodička, E. Holbrook, V.R. Ramakrishnan, M. Lechner, T. Hummel.

Author Affiliations:

*ENT department Asst Settelaghi- Varese University of Insubria Varese Italy

2nd ORL Department, Aristotle University, Thessaloniki, Greece

Jayant M. Pinto -Section of Otolaryngology-Head and Neck Surgery, The University of Chicago, Chicago, IL

AW. Fjaeldstad - Flavour Clinic, University Clinic for Flavour, Balance and Sleep, Department of Otorhinolaryngology, Regional Hospital Gødstrup, Denmark

Joaquim Mullol - Rhinology Unit & Smell Clinic, ENT Department, Hospital Clínic, IDIBAPS, Universitat de Barcelona, CIBERES. Barcelona, Catalonia, Spain.

Jan Vodička Department of Otorhinolaryngology and Head and Neck Surgery, Regional Hospital and University of Pardubice, Czech Republic; Department of Otorhinolaryngology and Head and Neck Surgery, Regional Hospital and Faculty of Heath Studies, University of Pardubice, Czech Republic

Jim Boardman, Fifth Sense, UK

Eric H. Holbrook, MD, MS, Department of Otolaryngology-Head and Neck Surgery, Massachusetts Eye and Ear, Harvard Medical School

Johannes Frasnelli, MD, Department of Anatomy, University of Québec in Trois-Rivières, Canada

Matt Lechner, MD, PhD, Division of Surgery and Interventional Science and UCL Cancer Institute, University College London and Barts Health NHS Trust, United Kingdom

Vijay R. Ramakrishnan, MD, Department of Otolaryngology-Head and Neck Surgery, Indiana University School of Medicine

Abstract:

Statement of problem: Evaluating the effectiveness of the management of Olfactory Dysfunction (OD) has been limited by a paucity of high-quality randomised and/or controlled trials. A major barrier is heterogeneity of outcomes in such studies. Core outcome sets (COS) –standardized sets of outcomes that should be measured/reported as determined by consensus—would help overcome this problem and facilitate future meta-analyses and/or systematic reviews (SRs). We set out to develop a COS for interventions for patients with OD.

Method(s) of Study: A long-list of potential outcomes was identified by a steering group utilising a literature review, thematic analysis of a wide range of stakeholders' views and systematic analysis of currently available Patient Reported Outcome Measures (PROMs). A subsequent e-Delphi process allowed patients and healthcare practitioners to individually rate the outcomes in terms of importance on a 9-point Likert scale.

Main results: After 2 rounds of the iterative eDelphi process, the initial outcomes were distilled down to a final COS including subjective questions (visual analogue scores, quantitative and qualitative), quality of life measures, psychophysical testing of smell, baseline psychophysical testing of taste, and presence of side effects along with the investigational medicine/device and patient's symptom log.

Principal conclusions: Inclusion of these core outcomes in future trials will increase the value of research on clinical interventions for OD. We include recommendations regarding the outcomes that should be measured, although future work will be required to further develop and revalidate existing outcome measures.

Key words: Olfactory Dysfunction, smell, core outcome set, effectiveness trial, outcome measurement

Introduction:

Olfactory dysfunction (OD) is a common yet under recognised and under treated condition[1]. Anosmia is thought to affect at least 5% of the general population but studies vary in prevalence and OD increases with age and can be as high as 20% in patients 60 years of age and older [2-5]; women are less commonly affected than men, albeit that they present to clinicians twice as much as men[6]. Apart from aging, common causes of OD include chronic rhinosinusitis (CRS) with and without nasal polyps, post-infectious olfactory dysfunction (PIOD) (including post-COVID-19), post-traumatic olfactory dysfunction (PTOD), allergic rhinitis, toxic exposures, neurological (e.g. Parkinson's, Alzheimer's), iatrogenic and idiopathic aetiologies[7 8]. Rarer causes of OD include olfactory bulb/ anterior skull base tumours, congenital aplasia, and olfactory cleft stenosis (OCS). With the onset of the global pandemic COVID-19, and nearly 60% of affected patients experiencing anosmia with the earlier variants, there has been an increase in the awareness of OD. Common sequelae of ODs include anxiety, depression, poor eating experience, isolation and malnutrition[9]. A recent exercise in priority setting for research in the UK has confirmed the clear need for more trials and interventions in this area[10].

To date there has been wide variability in studies and varied approaches to the topic across the globe. Multiple studies also have mixed aetiology groups and these factors have limited our ability to draw accurate conclusions which subsequently hinders the study of the impact of smell and taste disorders and treatment options[7]. Historically, studies in this field have used variable outcome measures, included participants with mixed aetiologies, and recruited samples sizes that are underpowered[11]. The Core Outcome Measures in Effectiveness Trials (COMET) Initiative which was launched in 2010 in the UK, and is supported by the National Institute of Health

Research, the Medical Research Council, the European Commission and the Seventh Framework Programme[12]. Although there is no specific methodology to generate a core outcome set, the majority follow a standard process of identifying existing knowledge by experts to develop a long list of outcomes, following an iterative Delphi process to develop consensus on key outcomes, leading to eventual global agreement across stakeholder groups.

Aim: The aim of our study is to develop a set of standard core outcome measures that can be used to study the effectiveness of treatment options in clinical trials of OD therapies. This will also better facilitate future systematic reviews and meta-analyses on the topic.

Materials and methods

COS development registration

Core Outcome Set (COS) development registration: The project was registered with the COMET Register, and the development process followed guidance issued by COMET. In particular, the minimum standards for COS development were met and the checklist for COS study reporting was followed. No ethical approval was required as opinions of health care professionals and patient representatives were included and no identifiable or individualised personal information was requested or used in this project. The setup used previously validated methods[12].

Defining scope

A participating group of Olfactologists (including ENT Surgeons with a special interest in olfactory disorders and clinical research scientists) and patient representatives was assembled for the Delphi process through personal invitation to members of the Clinical Olfactory Working Group (COWoG) by the senior author (CP). (COWoG – see website for details of membership:

https://www.uniklinikum-dresden.de/de/das-klinikum/kliniken-polikliniken-institute/hno/forschung/interdisziplinaeres-zentrum-fuer-riechen-und-schmecken/downloads-links/european-clinical-olfactory-working-group-ecowg)

Due to the high global variation and heterogeneous nature of previous studies, the group agreed for the need to undertake this process to include the most relevant outcome measures for use in interventional studies pertaining to smell and taste disorders. The COS is primarily aimed for use in clinical research, but the group agreed it could also be suitable for routine clinical care in specialist centres.

Stakeholder involvement

Both patient representatives, researchers and clinician experts in olfactory disorders were involved in every stage of COS development, including defining scope, developing the long list of outcome measures, the iterative Delphi process, review, and analysis of final results. Patient representatives were members of the public and patient involvement panel of the UK charity Fifth Sense (www.fifthsense.org.uk).

Delphi process

The first round of the two rounds of Delphi processes was held online in January 2022, the second was held online in March 2022. The timeline is depicted in figure 1.

Long-list development

An extensive list of potential core outcome measures was drawn up from the assembled group (Table 1). We invited the aforementioned participants to take the survey via Google Forms. There is no set number of participants for a Delphi process, and thus a pragmatic approach was taken. In the first round, each participant was asked to consider each outcome measure on a 9-point

Likert scale (figure 2) and also asked for additional suggestions. Scores of 7-9 were given for outcomes considered to be essential, scores of 4-6 given for outcomes thought to be optional and scores of 1-3 given for outcomes considered to be excluded. Individual responses were anonymous to other participants but not to the lead author. Responses were exported into an excel file and median outcome scores were calculated. At the end of the first cycle, the distribution of votes on each outcome measure was revealed to the group and discussed. Additional suggestions were discussed, and outcomes were amended/added by consensus. No outcomes were excluded at this stage.

Short-list development

The participants were then asked to complete the second Delphi cycle by completion of the survey via Google Forms. Participants scored using the same Likert scale as before, but with the knowledge of the previous set of results. The second cycle results were then calculated and discussed at the end of the second Delphi cycle to develop the Final Core Outcome Set.

Results

Delphi Cycle 1:

The first round Delphi process was held in January 2022. This included 25 participants in total. There were 19 healthcare and research professionals and 6 patient representatives. Amongst the survey responses, there was close agreement amongst healthcare and research professionals. In contrast, there were marked differences in responses from patient representatives. Clinical measures were rated highly by the clinicians. Specific quality of life measures was preferred by patient representatives (for example, SelfMOQ) compared to generalized measures (for example EQ-5D). Cost to healthcare system and cost incurred to patient was also rated higher by patient

representatives compared to health care professionals. From the long-list, nine items were regarded as essential to the core outcome set by all respondents. Table 1 shows the details of the long-list discussed. Table 2 shows the voting responses in both of the two Delphi cycles.

Delphi Cycle 2

The second round Delphi process was held in February 2022. This included 21 participants. There were 17 healthcare and research professionals and 4 patient representatives. There was a better understanding of outcome requirements and focus was on identifying inexpensive, easy to use, reliable, valid, standardised and globally recognisable measures. Many outcome measures in the list that were felt to be highly specific were considered for addition to extended / optional outcome measures list. One example of this was the Sinonasal Outcomes Test-22 (SNOT-22) score for interventional studies specifically pertaining to chronic rhinosinusitis (CRS) where only one specific question addresses OD; this measure was also previously included in the COMET initiative for CRS (CHROME)[13]. Any outliers were discussed and consensus was achieved.

Final Core Outcome Set

At the end of the two-stage Delphi process, outcome measures with a median score of 7 or more were taken as the final outcome measures to be included (see figure 3). This resulted in 5 key recommendations (including 4 outcome measures) that were considered essential to be measured in clinical trials of olfactory disorders include (See Table 3):

1. Visual Analogue Scores (quantitative and qualitative assessment of olfactory function)
2. Psychophysical smell testing (validated for the country and language of use): Sniffin' Sticks Test[14]/ University of Pennsylvania Smell Identification Test (UPSIT)[15]
3. Health-related quality of life (HRQoL) outcome measure:

a. Disease specific: Questionnaire of Olfactory Disorders (QOD)[16]

b. Generic: EQ-5D[17]

4. Patient symptom log (unspecified format)

The group also recommended taste measurement at baseline assessment using taste strips, not as a core outcome measure, but an essential measure to exclude any additional gustatory dysfunction. Table 4 lists the optional/extended list outcome measures that could be considered in specific studies where the OD or assessment of it, requires certain additional outcome measures to be included and resources are available to deliver them. For example, the APOLLO trial is a proof of concept study and has selected olfactory bulb volume (on MRI scans) as the primary outcome measure, with secondary outcomes including fMRI and DTI but has included the core outcome set[18]. Excluded outcomes are listed in table 5.

Discussion:

Key Results

The final COS has delineated a small number of outcome measures: a VAS, a validated psychophysical test, disease-specific and generic HRQoL measures and a patient log, that should provide clinician researchers globally with the means to standardise clinical trials in OD without great expense or the need for unwieldy specialist equipment. Researchers will have the option to use the extended list of core outcomes where appropriate for specific studies or where equipment and expertise are available. The COWoG also chose to include a baseline assessment of taste assessment, due to the common misperception between flavour and taste[19]. It was felt by the authors that these were important and essential elements in any trials for ODs but deliverable for researchers globally who should be able to include these outcomes without them

being prohibitive from a resource or economic perspective. Of course, the core set does not preclude researchers from additionally including outcomes from the extended list such as imaging modalities and other psychophysical tests; each trial design needs to consider an appropriate primary outcome measure for its purpose, but by including the ODs COS, allows for direct comparison across trials.

Limitations

A specific systematic review was not performed, however with access to an expert panel who represents active clinicians and researchers in the field of current research in the field, the group considered sufficient evidence to form the basis of the COMET process. Unfortunately, there was a 16% attrition rate from the first Delphi round meeting to the second, despite multiple reminders and due to the unavailability of panel members to attend the meeting. We opted for the benefits of an international group, but this entailed the complexity of scheduling the meetings. We also initially considered including a wider group of ENT specialists, but the presence of an expert panel and patient participation was considered adequate in providing specific expert input in an area of niche subspecialisation.

Interpretation

In comparison to the previous COS developed in the field of Rhinology for rhinosinusitis (CHROME)[13], this COS was at first glance a smaller list than the CHROME one. However, the CHROME domains were Patient Symptoms and QoL, Control of Disease, Impact on Daily Activity and Acceptability of Treatment and Side-Effects; the 7 listed outcomes shared many similarities such as HRQoL outcomes and assessment of treatment side-effects. Of course, researchers

running trials in CRS may in future choose to include both the rhinosinusitis COS and the OD COS where certain outcome measures will serve both needs across the two COSs.

In the field of smell and taste disorders, there is a lack of compelling evidence behind treatment options due to poorly designed studies, and thus there is a paucity of well-designed clinical trials to help guide clinicians in advice and treatment options for patients[8 11]. For example, when considering sample sizes, in 2015 Schopf et al. published a prospective controlled pilot study with less than 10 participants which is too small to infer clinical significance[20]. Similarly, Henkin et al. in 2017 published a prospective controlled study to assess the response to theophylline; not only did the study involve patients with mixed aetiologies but it also used a non-standardised smell test to report results[21]. A large number of similar studies identified from the COWoG consensus paper of post-infectious olfactory dysfunction[22] highlights the need for careful consideration of study design and research methodology in the future and a collective responsibility for groups such as COWoG to set a precedent for improving the quality of clinicals trials delivered for ODs in the future. This may include work to ensure adequate minimum clinically important differences (MCIDs) are available for selected outcome measures to ensure power calculations for primary outcome measures are appropriate[23].

Generalisability

The global standardisation of core outcome measures undertaken here can increase the strength of future systematic reviews and meta-analysis including the evidence from international consensus statements, for example the recent ICAR-Olfaction consensus statement by Patel et al[8]. The COWoG will promote dissemination of this COS through various media and platforms including conferences and seminars. It will also be available through the COMET website and

other professional social media channels/websites, for example Fifth Sense (www.fifthsense.org.uk; a patient charity based in the UK) and the Technical University of Dresden's Clinical Olfactory Working Group website (https://tinyurl.com/5cb7pmzn). This COS exercise will also provide the COWoG an opportunity to consider the most useful olfactory questionnaires and supporting global standardisation further. The COWoG will plan to revisit this exercise in 2027 so that any new outcome measures can be included as well as allowing for any changes in perception about the importance of the existing outcome measures.

Authorship contribution

Based on IJCME criteria, CP designed project, KK corresponded with contributing panel participants, arranged consensus meetings, executed the study and drafted the paper. All other authors offered their expert opinion via the Delphi process, performed oversight of the project, edited the draft, and approved the final manuscript.

Conflicts of interest:

- CP COIs outside this work: Grants from NIHR, Royal College of Surgeons, ESPRC, Sir Jules Thorn Trust; Honoraria/Fees from Stryker, GSK, Sanofi, Abbot, Olympus; Trustee of Fifth Sense
- JM COIs outside this work: Grants from AstraZeneca, Genentech, GSK, Viatris/MEDA Pharma, Novartis, Regeneron, Sanofi-Genzyme and Noucor/Uriach Group, consulting fees from Sanofi-Genzyme and Noucor /Uriach Group and attended speaker bureaus and/or advisory boards for AstraZeneca, Genentech, GSK, Glenmark, Menarini, Mitsubishi-Tanabe Pharma, MSD, Viatris/ MEDA Pharma, Novartis, Proctor & Gamble, Regeneron, Pharmaceuticals Inc., Sanofi-Genzyme, UCB Pharma and Noucor/Uriach Group

- TH COIs outside this work: Since 2018 TH did research together with and/or received funding from Sony, Stuttgart, Germany; Smell and Taste Lab, Geneva, Switzerland; Takasago, Paris, France; aspuraclip, Berlin, Germany; Bayer healthcare, Berlin, Germany; Baia Foods, Madrid, Spain, and Frequency Therapeutics, Farmington, CT, USA; Primavera, Oy-Mittelberg, Germany; Novartis, Nürnberg, Germany;

References:

1. Ball S, Boak D, Dixon J, et al. Barriers to effective health care for patients who have smell or taste disorders. *Clin Otolaryngol* 2021 doi: 10.1111/coa.13818
2. Vennemann MM, Hummel T, Berger K. The association between smoking and smell and taste impairment in the general population. *J Neurol* 2008;255(8):1121-6. doi: 10.1007/s00415-008-0807-9
3. Landis BN, Konnerth CG, Hummel T. A study on the frequency of olfactory dysfunction. *Laryngoscope* 2004;114(10):1764-9.
4. Bramerson A, Johansson L, Ek L, et al. Prevalence of olfactory dysfunction: the skovde population-based study. *Laryngoscope* 2004;114(4):733-7. doi: 10.1097/00005537-200404000-00026
5. Mullol J, Alobid I, Marino-Sanchez F, et al. Furthering the understanding of olfaction, prevalence of loss of smell and risk factors: a population-based survey (OLFACAT study). *BMJ Open* 2012;2(6) doi: 10.1136/bmjopen-2012-001256
6. Yang J, Pinto JM. The Epidemiology of Olfactory Disorders. *Current otorhinolaryngology reports* 2016;4(2):130-41. doi: 10.1007/s40136-016-0120-6 [published Online First: 2016/05/01]
7. Hummel T, Whitcroft KL, Andrews P, et al. Position paper on olfactory dysfunction. *Rhinol Suppl* 2017;54(26):1-30.
8. Patel ZM, Holbrook EH, Turner JH, et al. International consensus statement on allergy and rhinology: Olfaction. *International forum of allergy & rhinology* 2022;12(4):327-680. doi: 10.1002/alr.22929
9. Philpott CM, Boak D. The impact of olfactory disorders in the United kingdom. *Chem Senses* 2014;39(8):711-8. doi: 10.1093/chemse/bju043
10. Boak D, Philpott C. Smell and Taste Disorders: Top 10 Research Priorities: Fifth Sense, 2022.
11. Addison A, Philpott CM. A systematic review of therapeutic options for non-conductive olfactory dysfunction. *The Otorhinolaryngologist* 2018;11(2):61-71.
12. Gargon E. The COMET (Core Outcome Measures in Effectiveness Trials) Initiative. *Maturitas* 2016;91:91-2. doi: 10.1016/j.maturitas.2016.06.007 [published Online First: 20160609]
13. Hopkins C, Hettige R, Soni-Jaiswal A, et al. CHronic Rhinosinusitis Outcome MEasures (CHROME), developing a core outcome set for trials of interventions in chronic rhinosinusitis. *Rhinology journal* 2018;56:22-32. doi: 10.4193/Rhin17.247
14. Hummel T, Sekinger B, Wolf SR, et al. 'Sniffin' sticks': olfactory performance assessed by the combined testing of odor identification, odor discrimination and olfactory threshold. *Chem Senses* 1997;22(1):39-52.
15. Doty RL, Shaman P, Dann M. Development of the University of Pennsylvania Smell Identification Test: a standardized microencapsulated test of olfactory function. *Physiology and Behavior* 1984;32(3):489-502.
16. Langstaff L, Pradhan N, Clark A, et al. Validation of the olfactory disorders questionnaire for English-speaking patients with olfactory disorders. *Clin Otolaryngol* 2019;44(5):715-28. doi: 10.1111/coa.13351
17. Feng Y-S, Kohlmann T, Janssen MF, et al. Psychometric properties of the EQ-5D-5L: a systematic review of the literature. *Quality of Life Research* 2021;30(3):647-73. doi: 10.1007/s11136-020-02688-y
18. Roopa KK, Sara B, Saber S, et al. A Double Blinded Randomised Controlled Trial of Vitamin a Drops to Treat Post-viral Olfactory Loss: Study Protocol for a Proof-of-concept Study for Vitamin a Nasal Drops in Post-viral Olfactory Loss (Apollo). Research Square, 2022.
19. Doty RL. Psychophysical measurement of human olfactory function, including odorant mixture assessment. Handbook of olfaction and gustation: CRC Press 2003:408-51 %@ 0429213646.

20. Schopf V, Kollndorfer K, Pollak M, et al. Intranasal insulin influences the olfactory performance of patients with smell loss, dependent on the body mass index: A pilot study. *Rhinology* 2015;53(4):371-8. doi: 10.4193/Rhino15.065
21. Henkin RI, Velicu I, Schmidt L. An open-label controlled trial of theophylline for treatment of patients with hyposmia. *Am J Med Sci* 2009;337(6):396-406. doi: 10.1097/MAJ.0b013e3181914a97 [published Online First: 2009/04/11]
22. Addison AB, Wong B, Ahmed T, et al. Clinical Olfactory Working Group consensus statement on the treatment of postinfectious olfactory dysfunction. *J Allergy Clin Immunol* 2021;147(5):1704-19. doi: 10.1016/j.jaci.2020.12.641
23. Copay AG, Subach BR, Glassman SD, et al. Understanding the minimum clinically important difference: a review of concepts and methods. *Spine J* 2007;7(5):541-6. doi: 10.1016/j.spinee.2007.01.008 [published Online First: 20070402]
24. Frasnelli J, Hummel T. Olfactory dysfunction and daily life. *Eur Arch Otorhinolaryngol* 2005;262(3):231-5. doi: 10.1007/s00405-004-0796-y [published Online First: 2004/05/05]
25. Zou LQ, Linden L, Cuevas M, et al. Self-reported mini olfactory questionnaire (Self-MOQ): A simple and useful measurement for the screening of olfactory dysfunction. *Laryngoscope* 2020;130(12):E786-E90. doi: 10.1002/lary.28419 [published Online First: 20191120]
26. Hopkins C, Gillett S, Slack R, et al. Psychometric validity of the 22-item Sinonasal Outcome Test. *Clinical Otolaryngology* 2009;34:447-54. doi: 10.1111/j.1749-4486.2009.01995.x
27. Ware Jr. J, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care* 1996;34:220-33. [published Online First: 1996/03/01]
28. Ware Jr. JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30:473-83. [published Online First: 1992/06/11]
29. Brooks R, Rabin R, de Charro F. The Measurement and Valuation of Health Status using EQ-5D: A European Perspective. Dorderecht: Kluwer 2003.
30. Lildholdt T, Fogstrup J, Gammelgaard N, et al. Surgical versus medical treatment of nasal polyps. *Acta Otolaryngol* 1988;105:140-43.
31. Lund VJ, Kennedy DW. Staging for rhinosinusitis. *Otolaryngology - Head & Neck Surgery* 1997;117:S35-40. doi: 10.1016/S0194-5998(97)70005-6
32. Panagou P, Loukides S, Tsipra S, et al. Evaluation of nasal patency: comparison of patient and clinician assessments with rhinomanometry. *Acta Otolaryngol* 1998;118:847-51.
33. Grymer LF, Hilberg O, Pedersen OF, et al. Acoustic rhinometry: values from adults with subjective normal nasal patency. *Rhinology* 1991;29:35-47.
34. Vogt K, Jalowayski AA, Althaus W, et al. 4-Phase-Rhinomanometry (4PR)--basics and practice 2010. *Rhinol Suppl* 2010(21):1-50.
35. Kobal G, Hummel T, Sekinger B, et al. "Sniffin' sticks": screening of olfactory performance. *Rhinology* 1996;34:222-6.
36. Doty RL, Shaman P, Kimmelman CP, et al. University of Pennsylvania Smell Identification Test: a rapid quantitative olfactory function test for the clinic. *The Laryngoscope* 1984;94:176-8.
37. Cain WS, Goodspeed RB, Gent JF, et al. Evaluation of olfactory dysfunction in the Connecticut Chemosensory Clinical Research Center. *The Laryngoscope* 1988;98:83???88. doi: 10.1288/00005537-198801000-00017
38. Cardesín A, Alobid I, Benítez P, et al. Barcelona Smell Test - 24 (BAST-24): validation and smell characteristics in the healthy Spanish population. *Rhinology* 2006;44:83-9.
39. Rojas-Lechuga MJ, Ceballos JC, Valls-Mateus M, et al. The 8-Odorant Barcelona Olfactory Test (BOT-8): Validation of a New Test in the Spanish Population During the COVID-19 Pandemic. *J Investig*

353 *Allergol Clin Immunol* 2022;32(4):291-98. doi: 10.18176/jiaci.0824 [published Online First:
354 20220509]
355 40. Briner HR, Simmen D. Smell diskettes as screening test of olfaction. *Rhinology* 1999;37:145-48.
356 41. Croy I, Hoffmann H, Philpott C, et al. Retronasal testing of olfactory function: an investigation and
357 comparison in seven countries. *Eur Arch Otorhinolaryngol* 2014;271(5):1087-95. doi:
358 10.1007/s00405-013-2684-9
359 42. Renner B, Mueller CA, Dreier J, et al. The candy smell test: a new test for retronasal olfactory
360 performance. *Laryngoscope* 2009;119:487-95.
361 43. Landis BN, Welge-Luessen A, Brämerson A, et al. "Taste Strips" – A rapid, lateralized, gustatory
362 bedside identification test based on impregnated filter papers. *Journal of Neurology*
363 2009;256:242-48. doi: 10.1007/s00415-009-0088-y
364 44. Fjaeldstad A, Niklassen AS, Fernandes HM. Re-Test Reliability of Gustatory Testing and Introduction
365 of the Sensitive Taste-Drop-Test. *Chem Senses* 2018;43(5):341-46. doi: 10.1093/chemse/bjy019
366 45. Rombaux P, Guérit JM, Mouraux A. Lateralisation of intranasal trigeminal chemosensory event-
367 related potentials. *Neurophysiol Clin* 2008;38(1):23-30. doi: 10.1016/j.neucli.2007.12.002
368 [published Online First: 20080122]
369 46. Lund VJ, Mackay IS. Staging in rhinosinusitus. *Rhinology* 1993;31:183-4.
370 47. Hutson K, Kumaresan K, Johnstone L, et al. The use of MRI in a tertiary smell and taste clinic: Lessons
371 learned based on a retrospective analysis. *Clin Otolaryngol* 2022 doi: 10.1111/coa.13968
372 [published Online First: 20220802]
373 48. Rombaux P, Duprez T, Hummel T. Olfactory bulb volume in the clinical assessment of olfactory
374 dysfunction. *Rhinology* 2009;47:3-9.
375 49. Shiino A, Morita Y, Ito R, et al. [Functional MRI of the human brain responses to olfactory
376 stimulation]. *No Shinkei Geka* 1999;27(12):1105-10.
377 50. Theaud G, Houde JC, Boré A, et al. TractoFlow: A robust, efficient and reproducible diffusion MRI
378 pipeline leveraging Nextflow & Singularity. *Neuroimage* 2020;218:116889. doi:
379 10.1016/j.neuroimage.2020.116889 [published Online First: 2020/05/21]
380 51. Rombaux P, Huart C, Mouraux A. Assessment of chemosensory function using
381 electroencephalographic techniques. *Rhinology* 2012;50:13-21. doi: 10.4193/Rhino11.126
382 52. Rombaux P, Mouraux A, Keller T, et al. Trigeminal event-related potentials in patients with olfactory
383 dysfunction. *Rhinology* 2008;46:170-4.
384 53. Hummel T, Stupka G, Haehner A, et al. Olfactory training changes electrophysiological responses at
385 the level of the olfactory epithelium. *Rhinology* 2018;56(4):330-35. doi: 10.4193/Rhin17.163
386 [published Online First: 2018/08/05]
387 54. Kobal G. Gustatory evoked potentials in man. *Electroencephalography and clinical neurophysiology*
388 1985;62:449-54.
389 55. Jafek BW, Murrow B, Michaels R, et al. Biopsies of human olfactory epithelium. *Chem Senses*
390 2002;27(7):623-8. doi: 10.1093/chemse/27.7.623
391 56. Sharma A, Kumar R, Aier I, et al. Sense of Smell: Structural, Functional, Mechanistic Advancements
392 and Challenges in Human Olfactory Research. *Curr Neuropharmacol* 2019;17(9):891-911. doi:
393 10.2174/1570159x17666181206095626
394 57. Horigome K, Pryor JC, Bullock ED, et al. Mediator release from mast cells by nerve growth factor.
395 Neurotrophin specificity and receptor mediation. *J Biol Chem* 1993;268:14881-87.

396

397

398 Table 1: List of items included in the long-list

Category	Outcome measure
Subjective questions	Qualitative VAS (0-10cm) Quantitative VAS (0-10cm)
Quality of life	Olfactory Disorders Questionnaire[24] Self-reported Mini Olfactory Questionnaire (SelfMOQ)[25] SNOT-22[26] SF-12[27] SF-36[28] EQ-5D[29]
Rhinological	Nasal endoscopy plus scoring (Lildholdt polyp score, Lund-Kennedy score)[30 31] Peak Nasal inspiratory flow[32] Acoustic rhinometry[33] Other airflow measurements (e.g. rhinomanometry)[34]
Psychophysical (not an exhaustive list)	Sniffin' Sticks[35] UPSIT (University of Pennsylvania smell identification test)[36] CCCRCT (Connecticut Chemosensory Clinical Research Center Test)[37] Barcelona Smell Test (BAST-24)[38] BOT-8[39] Smell Diskettes[40] Retronasal testing – taste powders[41] Retronasal testing – candy smell test[42] Taste sprays Taste strips[43] Taste Drop Test[44] Trigeminal lateralisation task[45]
Radiology	CT (Computerised Tomography) scan (plus scoring, e.g.: Lund MacKay score)[46] MRI scan[47] MRI Volumetric measurements[48] Functional MRI[49] Diffusion weighted MRI[50]
Electrophysiological	OERPs (Olfactory Event-Related Potential)[51] Trigeminal ERPs (Event-related Potential)[52] Electro-olfactogram[53] GERPs (Gustatory Event-Related Potential) [54]
Pathophysiological	Olfactory biopsies/brushing[55] Olfactory binding protein[56] Brain derived neurotrophic factor[57]

Acceptability of treatment and compliance	Clinical records: History and Examination findings Presence of side effects (medication related) to the investigational medicinal product Patient diary Weight of medicine containers returned at follow up visits Cost incurred by patient Cost to healthcare system

399

400

Table 2: Results from iterative Delphi process (Cycle 1 and 2): Median scores for the group as a whole are represented for each cycle against each outcome measure voted on. Red (scores 1-3) indicates an outcome to be excluded, transitioning through yellow (scores 4-6) for outcomes considered optional, to green (scores 7-9) indicating an outcome to be included.

List of considered Core Outcome Measures	Delphi 1	Delphi 2
Visual analogue score (qualitative)	7.5	9
Visual analogue score (quantitative)	8	9
Questionnaire of Olfactory Disorders (QOD)	7	8
SNOT-22	6.5	5
SF-12	5	5
SF-36	5	4
EQ-5D	4	4
SelfMOQ	5	3
Nasal endoscopy plus scoring (Lildholdt polyp score, Lund Kennedy score)	8	9
Peak nasal inspiratory flow	5	5
Acoustic rhinometry	3	2
Other airflow measurements (e.g. rhinomanometry)	5	3
Sniffin' Sticks	9	9
UPSIT (University of Pennsylvania smell identification test)	8	7
CCCRCT (Connecticut Chemosensory Clinical Research Center Test)	6	6
Smell diskettes	4	5
Retronasal testing - taste powders	5	5
Retronasal testing - candy smell test	5	5
Taste sprays	6	7
Taste strips	7	7
Trigeminal lateralization task	5	5
CT scan (plus scoring, e.g., Lund MacKay score)	5	5
MRI scan	6	5
MRI: Volumetric measurements	5	5
Functional MRI	4.5	3

Diffusion weighted MRI	4	3
OERPs (Olfactory Event-Related Potential)	5	4
Trigeminal ERPs (Event-Related Potential)	5	4
Electro-olfactogram	4.5	3
GERPs (Gustatory Event-Related Potential)	4.5	2
Olfactory biopsies	4	3
Olfactory binding protein	3.5	2
Brain derived neurotrophic factor	3	2
Clinical records: History and Examination findings	9	9
Presence of side effects (medication related) to the investigational medicinal product/ device	9	9
Patient diary	6	7
Weight of medicine containers returned at follow up visits	5	4
Cost incurred by patient	5.5	5
Cost to healthcare system	6	6

406

407

408 Table 3: Finalised Core Outcome Set:

Key COS Domains	Choice of Outcome Measures
Patient Reported Outcome Measures	Quantitative and Qualitative Visual Analogue Score
Quality of life measures	Questionnaire of Olfactory Disorders Questionnaire (QODQ), EQ-5D
Psychophysical testing	Sniffin' Sticks Smell Test or UPSIT
Presence of side effects (medication related) to the investigational medicinal product/ device	Patient diary/ Symptom log
Baseline gustatory function assessment (not an outcome measure)	*Taste strips*

409

410 Table 4: Extended list/ Optional outcome measures:

Recommendations for optional Outcome Measures/ Extended List	
SNOT22 (Sinonasal Outcomes Test 22)	For studies in Chronic Rhinosinusitis (CRS) patients
Nasal endoscopy plus various scoring measures (Liltholdt score and Lund-Kennedy score)	For CRS patients
Peak nasal inspiratory flow (PNIF)	e.g. GM instruments PNIF meter
Other psychophysical tests	Smell diskettes or other (newer) smell tests
Retronasal testing	taste powders, candy smell test
Taste sprays	Custom made
Trigeminal lateralization task	e.g. CO_2 stimulation
Radiological imaging	CT, MRI (fMRI, dwMRI)
Electrophysiological testing	OERPs
Compliance measures to intervention	Weight of medicine
Health economic measures	Cost incurred to patient; Cost incurred to healthcare system, SF-12

411

412 Table 5: Outcome measure excluded from iterative Delphi process:

Outcome measures excluded:
SelfMOQ
fMRI
dwMRI
Electro-olfactogram
GERPs
Trigeminal ERPs
Olfactory binding protein
BDNF

413

414

Figure 1: Delphi timeline showing process of development of the COS. COWoG = clinical olfactory working group, HCP = health care practitioner, PPI = patient/lay representative

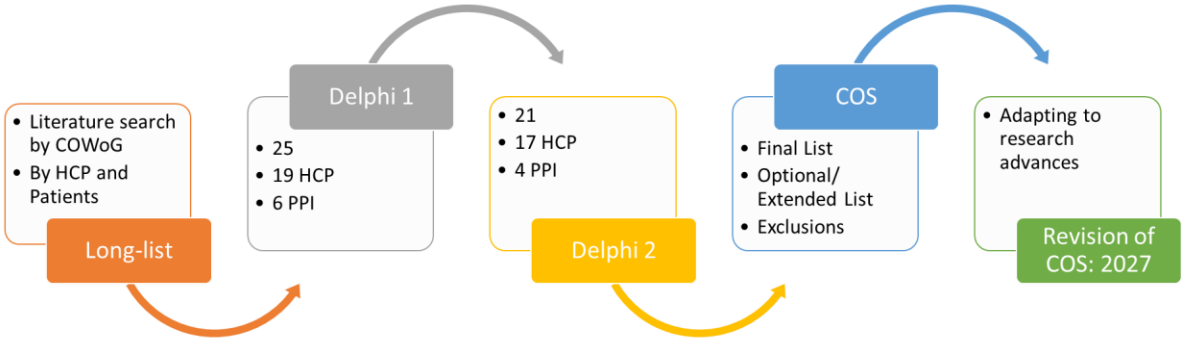

Figure 2: Nine-point Likert scale indicating how each score represented each participant's view of whether or not the outcome measure should be included.

Strongly Disagree	Disagree	Moderately Disagree	Mildly Disagree	Undecided	Mildly Agree	Moderately Agree	Agree	Strongly Agree
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)

Figure 3: Median responses for each considered outcome measure; those scoring 7 or more at the second Delphi were included.