WP2 / WP6 – Core Outcome Set Project

DELPHI - Core Outcome Set (COS) definition in

Waldenstrom's Macroglobulinaemia (WM)

December 1st, 2021

INDEX

A. INTRODUCTION

The HARMONY Alliance is a public-private European Network established in 2017, which includes 53 partners and 43 associated members from 17 countries, including 9 pharmaceutical companies and 9 Patient Umbrella Organizations. One of HARMONY's objectives is to use Big Data to improve understanding and treatment of hematological malignancies (HM) (1). HARMONY Plus is a new public-private partnership within the HARMONY Alliance, launched at 6 October 2020. One of Harmony Plus objectives is to expand the scope of the HARMONY Alliance to cover remaining HM not included in the HARMONY project (2). Just like the previous HARMONY project one work package within HARMONY Plus is focused on defining outcomes sets for further HMs and one outcome set applicable for all HMs. In accordance, this study will be performed to define the core outcome set (COS) in Morbus Waldenström (MW) or Waldenström's Macroglobulinaemia (WM), one out of four hematological malignancies predefined in HARMONY Plus.

Waldenström macroglobulinaemia is an indolent B-cell lymphoma with clearly defined criteria for diagnosis, initiation of therapy, and response (3). WM is a rare disease of the elderly that accounts for 1%–2% of non-Hodgkin lymphomas. The diagnosis of WM is based on the histopathological confirmation of bone marrow infiltration by lymphoplasmacytic cells/lymphoplasmacytic lymphoma (LPL) and the detection of any amount of monoclonal immunoglobulin M (IgM) protein (4). Differentiation from other entities, such as monoclonal gammopathy of undetermined significance (MGUS), is essential. Risk assessment is currently based on the international scoring system for Waldenström macroglobulinemia (IPSSWM score) (5, 6). Various treatment options such as are available, the choice of therapy is based on the need for rapid disease control, presence of specific disease complications, and patient's age (7).

Generally valid recommendations of outcomes that should be measured are still missing.

Unfortunately, the ability to compare clinical trials is limited due to differences in their measured outcomes. This lack of standardization relates to the current lack of a COS that can be utilized to guide outcomes selection and harmonization in WM in current and future trials. For example, measurement of long-term side effects and their influence on the patients' quality of life has not yet been assessed in most of clinical trials.

A COS is a minimum set of outcomes developed by consensus, and a minimum set of outcomes is a reference point and provides the minimum outcomes that should be collected in further clinical trials on a given condition. It is common to develop a COS by consensus by using multi-stakeholder consensus-based Delphi methodology. Use of a COS improves the comparability of clinical trials or other research in real world settings, improves consistency of reporting, reduces selective reporting bias and ensures that appropriate outcomes valued by a range of stakeholders are measured. COS can be incorporated into clinical guidelines and improve the clinical practice and patient outcomes and management.

Key stakeholders who are dedicated to provide their expert feedback are selected based on their skills and experience relevant to the disease or project. The stakeholders include health service users, health service practitioners, researchers, regulators, drug developers, patients and patient advocates. Participants of all stakeholder groups were in particular recruited from members of the HARMONY work packages, but also participants outside the HARMONY Alliance are welcome to take part of the Delphi survey within their stakeholder group.

In order to ensure that the defined COS is acceptable for each stakeholder group it is important to include as many stakeholders' groups as possible in particular patients and patient advocates to increase the influence of patient groups for the definition of outcomes. An additional category is included in the analysis of the Delphi survey, called "patient important". This category will be used in the final analysis to mark a specific outcome as patient important. It is recommended to discuss these specific outcomes separately in the final consensus meeting.

B. PROJECT GOALS

The aim of this project is therefore to define a COS for WM agreed by consensus of all stakeholder groups and to define standardized outcomes to be measured in future clinical trials and observational studies throughout Europe.

The protocol has been written following the COS-STAP recommendations (8).

C. METHODS

The development of the COS will follow COMET recommendations from the international COS-STAD study (8,9).

To achieve consensus from different stakeholder groups the Delphi method will be used. The Delphi instrument used is an online tool, DelphiManager, provided by the COMET Initiative (10). A more detailed description of the methodology can be found in section D. Recruitment of participants mainly takes place from members of the HARMONY Alliance.

Participants

1. Patients

In this Delphi survey patients equal or older than 18 years with WM can participate. Different subtypes of WM are equally included, regardless of previous treatments including stem cell transplantation. Patients treated as outpatients are included as well as patients treated in hospital settings. Due to the use of English for the Delphi survey, participation is limited to patients understanding English.

2. Clinicians and Clinical researchers

Every clinician within or outside the HARMONY Alliance with experiences in WM can take part in the survey.

3. Drug developers

Participants have been recruited from stakeholder organizations that are members of HARMONY Plus, including European Federation of Pharmaceutical Industries and Associations (EFPIA) member companies.

4. Regulators

Recruitment of participants will be performed within the HARMONY Alliance with support of Work Package 6.

Data protection

The personal data of participants (name, home country and email address) will be stored only for the duration of the survey on a secure server provided by the DelphiManager. After completion of the survey all data will be deleted.

By registering, all participants provide consent to the terms of the Delphi survey and they agree to the use of their data in the way described in the survey protocol.

Selection of the outcome list for WM

The empirical basis for identifying a list of preliminary WM outcomes for the Delphi study so far has been threefold a two-step process:

First – A literature research was conducted in the COMET database to get an overview of the outcomes already used in existing clinical trials (11). The primary outcomes list was generated by extracting outcomes from the published literature (3-7).

Second – in order to include the patients' perspective, patient advocates and people who have or have had WM were invited to complement the preliminary list of outcomes by including additional outcomes and revise the list in accordance with their comments. In addition, a specific literature research for patient-reported outcomes in WM-patients was performed and included in the preliminary list (12).

D. DELPHI PROCESS

The preliminary WM outcome list created after the process described above (Annex 1), will be used in the Delphi survey in a representative pool of stakeholders to agree in a pre-defined and iterative process on a COS for WM.

The Delphi survey will include two rounds. In each round, the stakeholders will be asked to rate the importance of each outcome based on their personal experiences. Each outcome will be ranked into three categories (1-3 "not important", 4-6 "important but not critical" and 7-9 "critical") using a Likert scale of 1 to 9. After the completion of the first round of the Delphi survey no new participant will be invited.

Based on the experience of the previous harmony surveys, the surveys planned now will be held as a so-called "hackathon".
For this purpose, a virtual meeting will take place on at least two days - this is also due to the current pandemic situation.
At these meetings, the surveys will be conducted in parallel by all participants. A major advantage of this is that any questions that arise can be asked and answered directly and, if necessary, support can be offered.

Within the questionnaire, outcomes will be grouped into domains so similar or related outcomes can be viewed and rated together. Each outcome will be described in plain language. Plain language descriptions are used from lists provided by patient advocates and also from native speakers with medical background.

When registering, participants will be asked which stakeholder group he/she belongs to. Once the individual participant has completed the first ranking round, he/she will also be able to provide additional feedback, by suggesting additional outcome parameters, which might be added within the

subsequent Delphi rounds. This additional outcome will be added to the following Delphi rounds when two or more participants proposed this outcome to be included.

After each round, all participants will be provided with their own answers and an anonymized summary of the other participants' answers across all different stakeholder groups, in terms of the percentage scoring each of 1 to 9 on a particular outcome. Thereby feedback is provided from all stakeholder groups separately.

This allows the participants to revise their answers during the next round of the Delphi survey by taking the previous round's results into account. No outcome will be dropped out, so the participants can revise their initial ranking. The range of answers should decrease from round to round and a consensus opinion result, a core outcome set is defined. The process is stopped after pre-defined consensus criteria as described below.

After the final round a face-to-face consensus meeting will take place to finally discuss the results and to reaffirm the defined COS.

It will be important that as many participants as possible complete every round of the Delphi survey to ensure robust results of high representativeness.
The rate of non-response after the Delphi rounds, so called attrition is often highly variable. The attrition rate described over different Delphi studies varies from 0% to 20%. There is no recommendation regarding attrition rates, however an acceptable response rate would be 80%. To increase the response rates personalized email reminders will be sent out.
Attrition bias may occur if participants give no response to subsequent rounds of survey. Little evidence is available regarding the extent to which attrition bias influences the Delphi result.

E. RESULTS AND ANALYSIS

To reduce potential bias in the interpretation of the results a clear definition of consensus is crucial. There are three categories of consensus:

1. **Consensus in**
   70 % or more respondents over all the respondents (clinicians, EFPIA members, regulators/HTA, patients and patient advocates) scored the outcome as critically important (7-9) AND 15% or fewer rate the outcome as limited important (1-3)

2. **Consensus out**
   70 % or more of all the respondents (clinicians, EFPIA members, regulators/HTA, patients and patient advocates) scored the outcome as limited important (1-3)
   AND 15 % or fewer rate the outcome as critically important (7-9)

3. **No consensus**

   Outcomes that do not achieve a consensus through the several rounds in the Delphi survey.

After completing the last Delphi round, each participant will be asked about willingness to participate in a final meeting, representatives from all stakeholder groups will be part of this meeting.

The analysis of the Delphi study described in this protocol will use descriptive statistics. The results for each of the Delphi rounds, for each outcome and for each stakeholder group, will be presented in frequency tables. Quantitative analysis of the Delphi survey include calculations of i) percentage of panel's response rates and ii) percentages of responses in each of the three importance categories (1-3:" not important", 4-6: "important but not critical" and 7-9: "critical" based on 9-point Likert scale) for each outcome.

The data will be also displayed graphically, e.g., using histograms, for each stakeholder group and for each outcome. The plots will be reproduced for each round to further visualize the stability of the panel's opinion.

The analysis of the Delphi study will be performed using the R statistical software version 3.5.2. As mentioned above the exploratory analysis of the outcomes considered as important for patients will be analyzed as following: The median Likert score for the patient group at the end of each round will be calculated and those outcomes achieving a median of greater or equal to 7 (≥7) will be considered as important to patients.

F. STRENGTH & LIMITATIONS

As mentioned above different stakeholder groups take part in the Delphi survey. To ensure the impact of the highly important patient involvement in this process, a further specific category was added, called “patient important”. Thereby outcomes with a special interest for patients can be marked and emphasized in analysis.

The language used in the Delphi survey is English. This limits the group of people to participate in the Delphi to persons who do speak English. This might introduce a bias with regard to the countries participating in the Delphi, with e.g., a potential overrepresentation of English-speaking countries. While it was considered to translate the questionnaires into other European languages, this could pose additional problems and might introduce a different bias, e.g., depending on quality of the translations or depending on the number of participants per language, to name only a few.

To date, there is no recommendation found in literature regarding the number of participants to include in a Delphi survey. For certain stakeholder groups, for example for regulators it may be hard to recruit a large number of participants, which may lead to an imbalance of group size. With providing summarized results for each stakeholder group separately, the effect of inequitable distribution of group size is minimized, as described by COMET (13).

G. OUTLOOK

The anticipated way of developing the COS ensures that clinicians, industry, health authorities, as well as patients and patient advocates are involved in each stage of the development. In addition, the Delphi survey helps to make sure that the COS represents the priorities of all stakeholders. Ultimately, utilization of the COS will improve the relevance of trial endpoints to all stakeholders. Furthermore, it will increase the capacity for data synthesis between different trials.

In parallel to the completion of the Delphi survey in WM, it is intended to start Delphi surveys to define a COS for CML, Hl included in HARMONY Plus

Finally, based on the results of the COS definition for each hematological malignancies included in HARMONY and HARMONY Plus a standardized COS applicable for all HMs will be created.

ANNEX 1 | PRELIMINARY OUTCOME LIST FOR WM

Name	HelpText	DomainName	DomainName - simplified
Sensory neuropathy	Problems involving damage to the peripheral nerves (those that connect the limbs and organs to the central nervous sysem and control sensation, movement and coordination) or symptoms caused by those issues, including numbness, tingling or burning sensations, increased sensivity to touch, weakness or dysfunction especially of extremities	PRO / HR-QoL - general - non-clinical	PRO
Pain	Unpleasant physical sensation, including aching joints, which may vary in intensity from mild discomfort to pain that limits activities of daily life, limits self care and/or requires medication or hospitalisation. Medication may be necessary	PRO / HR-QoL - general - non-clinical	PRO
Diarrhea	Passing looser stools (poo) or passing stools more often than is normal for you	PRO / HR-QoL - general - non-clinical	PRO
Constipation	Having difficulty passing stools (poo), which may be small and hard	PRO / HR-QoL - general - non-clinical	PRO
Nausea	Feeling or being sick, which may lead to impact on intake of food and/or fluids and/or normal activities	PRO / HR-QoL - general - non-clinical	PRO
Changes in taste and smell	Loss of the senses of smell and taste, including the reduced ability to smell or taste specific substances, for instance, sweet, sour, bitter or salty	PRO / HR-QoL - general - non-clinical	PRO
Anorexia	Loss of appetite, which may lead to weight loss and malnutrition	PRO / HR-QoL - general - non-clinical	PRO
Fatigue	Significant or persistant tiredness that's not proportional to recent activity	PRO / HR-QoL - general - non-clinical	PRO

Shortness of breath (Dyspnoea)	Shortness of breath or respiratory problems, which may happen at rest and may limit activities of daily living or self care, and may require treatment	PRO / HR-QoL - general - non-clinical	PRO
Change in sexual function	Such as changes in sexual desire, sexual dysfunction, erectile dysfunction, difficulties reaching orgasm, vaginal dryness in women, other genital changes that lead to pain during sexual activity, difficulty feeling arousal and pleasure during sex	PRO / HR-QoL - general - non-clinical	PRO
Infertility	Inability to get pregnant or to produce healthy sperms	PRO / HR-QoL - general - non-clinical	PRO
Hair loss	Alopecia or baldness, loss of hair from part of the head or body	PRO / HR-QoL - general - non-clinical	PRO
Sleep changes	Finding it difficult to get to sleep or to stay asleep	PRO / HR-QoL - general - non-clinical	PRO
Anxiety	Feelings of constant worry, or deep concern or uneasy about uncertainties	PRO / HR-QoL - general - non-clinical	PRO
Depression	Feelings of severe sadness and unhappiness, often with decreased energy, constant feelings of guilt, doubt or self-blame, worthlessness and hopelessness	PRO / HR-QoL - general - non-clinical	PRO
thrombosis / thromboembolism	a blood clot that forms a vein, results in pain or embolisms, thrombosis broke up and travelled e.g. to the lung, that results in dyspnea or ultimatively death	PRO / HR-QoL - general - non-clinical	PRO
bleeding	blood loss	PRO / HR-QoL - general - non-clinical	PRO
hyperviscosity	because of too many components in the blood, the flow properties change and organ damage can occur	PRO / HR-QoL - general - non-clinical	PRO

Psychosocial function	Problems with mental processes of perception, memory, judgment, reasoning or thinking with an effect on relationships with partner, family and friends including ability to join in with social activities	PRO / HR-QoL - general - non-clinical	PRO
Physical function	The effect of M. Waldenstroem or its treatment on day to day physical activities; for example, walking, climbing stairs, driving	PRO / HR-QoL - PRO domains	PRO
Role function	The effect of M. Waldenstroem or its treatment on your role; for example, ability to look after children or to work or earn money	PRO / HR-QoL - PRO domains	PRO
Financial toxicity	Financial losses because of co-payment for medical treatment, and if a patient was working before disease diagnosis or progression, loss of salary during sick leave, which may include leave taken by a carer	PRO / HR-QoL - PRO domains	PRO
Eating and drinking	The effect of M. Waldenstroem or its treatment on eating and drinking	PRO / HR-QoL - PRO domains	PRO
Cost of treatment	Money which must be spend on M. Waldenstroem treatment	Health resource utilization - resource use	resource use
Emergency Unit admissions	Emergency or unplanned hospital treatment is necessary	Health resource utilization - resource use	resource use
Intensive care admissions	Requirement for treatment on an intensive care ward due to serious or life threatening disease progression or side-effects	Health resource utilization - resource use	resource use
Outpatient visits	Treatment or diagnostic visits in hospital without spending a night there	Health resource utilization - resource use	resource use
Need of caregiver assistance	Requirement for assistance given by caregiver (who could be a family member, friend or a	Health resource	resource use

	professional care giver) in or outside the hospital	utilization - resource use	
Complete Response - CR (complete remission)	M. Waldenstroem gets better, resulting in no evidence of plasma cells in tissues or bone marrow and negative immunfixation of serum and urine	Clinical outcome - Event type	type of event
Partial Response - PR (partial remission)	M. Waldenstroem gets better, with a substantial reduction of measuable sites or paraprotein burden compared to levels before treatment, but not enough to qualify as CR	Clinical outcome - Event type	type of event
Response - Stable disease (SD)	M. Waldenstroem stays the same after treatment. The cancer is not getting better or worse	Clinical outcome - Event type	type of event
Very good partial Response (VGPR)	Good response that fits special criteria.	Clinical outcome - Event type	type of event
Minimal response (MR)	Treatment was not that effective, but the M. Waldenstroem showed a response.	Clinical outcome - Event type	type of event
Relapse - Clinical relapse	Symptomatic return of M. Waldenstroem after a patient initially responds well to treatment	Clinical outcome - Event type	type of event
Cause of death	Death for any reason, whether related to M. Waldenstroem or not. This records the specific reason for death, not the time until death	Clinical outcome - Event type	type of event
Progressive disease (PD)	Worsening of a patient's M. Waldenstroem defined by a set of specific criteria	Clinical outcome - Event type	type of event
Overall survival (OS)	Length of time that a patient remains alive from either the date of diagnosis or the start of treatment for the M. Waldenstroem	Clinical outcome - Time to event	time to event
Progression free survival (PFS)	Time until someone's M. Waldenstroem either gets worse or they die from any cause	Clinical outcome - Time to event	time to event
Event free survival (EFS)	Time until someone's M. Waldenstroem either gets worse, they die from any cause or they	Clinical outcome - Time to event	time to event

	stop their treatment because of side-effects		
Duration of response (DOR)	Length of time from responding positively to a treatment to the M. Waldenstroem starting to recur / to get worse	Clinical outcome - Time to event	time to event
Time to progression (TTP)	Time until someone's M. Waldenstroem recurs / gets worse (excluding death)	Clinical outcome - Time to event	time to event
Treatment free intervall (TFI)	Time from the end of the treatment until the next therapy is needed	Clinical outcome - Time to event	time to event
Time to next treatment TTNT	Time from the end of primary treatment until the institution of the next therapy	Clinical outcome - Time to event	time to event
Infections	How often and how bad a patient gets sick or picks up a bacterial, viral or fungal infection, that needs antibacterial or antifungal treatment. Number of bacterial, viral or fungal infections, that needs antibacterial or antifungal treatment	Clinical outcome - clinical parameter	clinical parameter
Transfusions indepence	Need of red cell or platelet transfusions	Clinical outcome - clinical parameter	clinical parameter
Minimal residual disease (MRD) molecular	The level of M. Waldenstroem that can be detected as measured by using a DNA sequencing technique	Clinical outcome - MRD	clinical parameter
Minimal residual disease (MRD) imaging	The level of M. Waldenstroem that can be detected as measured by using a CT or PET-CT scan	Clinical outcome - MRD	clinical parameter
Minimal residual disease (MRD) flowcytometric	The level of M. Waldenstroem that can be detected as measured by using flowcytometry	Clinical outcome - MRD	clinical parameter
AEs (adverse events) and SAEs (serious adverse event)	A negative event or side-effect that happens during or after treatment, a clinical decision classified according to the latest "Common Terminology Criteria for Adverse Events", a descriptive terminology of adverse events.	Safety outcome - AE / Toxicity	safety outcome

	For each adverse event there is a grading for severity		
Medication adherence	Patients take their medication as prescribed by the doctor	Safety outcome - AE / Toxicity	safety outcome
Discontinuation of treatment	Patient decides to stop treatment themselves or under the direction of his/her doctor for any reason other than finishing a course of treatment	Safety outcome - AE / Toxicity	safety outcome
Hematological toxicity	Side-effects that cause changes in the blood or number of blood cells (e.g. anemia, leukopenia, thrombocytopenia, among others)	Safety outcome - AE / Toxicity	safety outcome
Non-Hematological toxicity	Side-effects that cause changes anywhere other than in the blood, e.g. nausea, neuropathy, mucositis, renal or liver failure, infections	Safety outcome - AE / Toxicity	safety outcome
neurological adverse events	Side-effects that causes neurological symptoms	Safety outcome - AE / Toxicity	safety outcome
Tolerability related outcomes	Measurement of how well patients are able to manage side-effects and whether they need to reduce dose or stop treatment as a result	Safety outcome - AE / Toxicity	safety outcome
Second primary malignancies (SPM)	A new cancer occurring in someone who has had a cancer in the past. It is different to recurrence, which is where the original cancer has returned	Safety outcome - AE / Toxicity	safety outcome

ANNEX 2 | REFERENCES

(1) **https://www.harmony-alliance.eu**

(2) HARMONY PLUS - HARMONY Alliance (harmony-alliance.eu)

(3) Castillo JJ et al. Consensus treatment recommendations from the tenth International Workshop for Waldenström Macroglobulinaemia. Lancet Hematol. 2020 7(11)

(4) Kastritis E et al. Waldenström's macroglobulinaemia: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol 2018. 29(4)

(5) Morel P et al. International prognostic scoring system for Waldenstrom macroglobulinemia. Blood. 2009 113

(6) Advani P et al. Updates in prognostication and treatment of Waldenstrom's macroglobulinemia. Hematol Oncol Stem Cell Ther. 2019. 12(4)

(7) Dimopoulos MA, Kastritis E: How I treat Waldenstrom macroglobulinemia. Blood. 2019. 134(23)

(8) Kirkham JJ et al. Core Outcome Set-STAndardised Protocol Items: the COS-STAP Statement. Trials 2019 20(116)

(9) Kirkham JJ, Davis K, Altman DG, Blazeby JM, Clarke M, et al. (2017) Core Outcome Set-STAndards for Development: The COS-STAD recommendations. PLOS Medicine 14(11)

(10) COMET Initiative | Home (comet-initiative.org)

(11) Research COMET Initiative | Home (comet-initiative.org) on 31.12.2021

(12) Tohidi-Esfahani et al. Whimsical (Waldenstrom's Macroglobulinemia Study Involving Cart-WheeL): A Global Patient-Derived Data Registry Capturing Treatment, Quality of Life and COVID-19 Data. Blood 2020. 136(1)

(13) Williamson PR, et al., The COMET Handbook: version 1.0. Trials, 2017. 18 (Suppl 3): p. 280

The HARMONY Alliance is funded through the Innovative Medicines Initiative (IMI), Europe's largest public-private initiative aiming to speed up the development of better and safer medicines for patients. HARMONY has received funding from IMI 2 Joint Undertaking and is listed under grant agreement No. 116026. This Joint Undertaking receives support from the European Union's Horizon 2020 Research and Innovation Programme and the European Federation of Pharmaceutical Industries and Associations (EFPIA). IMI supports collaborative research projects and builds networks of industrial and academic experts to boost pharmaceutical innovation in Europe.

www.harmony-alliance.eu

HARMONY Communications Office
European Hematology Associations (EHA), The Hague, The Netherlands
- communications@harmony-alliance.eu

HARMONY Coordination Office
Institute of Biomedical Research of Salamanca (IBSAL), Salamanca, Spain
- harmonyoffice@ibsal.es