

Feasibility and design for the preparation of a randomised control trial-based evaluation:

Remote versus face-to-face asylum personal interviews

Final Report

Prepared by RAND Europe for the European Union Agency for Asylum (EUAA)

The sole responsibility for this report lies with the author. The EUAA is not responsible for any use that may be made of the information contained therein.

EUAA/EVAL/2022/16/FR; Final

March 2023

TABLE OF CONTENTS

Contents

Executive Summary	vi
1. Introduction	1
1.1. Background to the personal interview and this study	1
1.2. Objectives of the Feasibility and Design Study	2
1.3. Methodological approach and the data collection that has informed this Study	4
1.4. Study strengths and limitations.....	4
1.5. The aim and structure of this Final Report	5
2. Feasibility of a RCT.....	6
2.1. Is it feasible to define the 'intervention' and implement it with fidelity in EU asylum determination contexts?.....	7
2.2. Are there feasible options for randomisation?.....	14
2.3. Is it feasible to define clear primary and secondary research questions and to collect data to answer those questions?	17
2.4. Is the number of interviews conducted likely to generate a sufficient sample size within a reasonable timeframe?	23
2.5. Is it feasible to implement a RCT ethically?.....	24
2.6. Is it feasible to implement a RCT in social and political contexts?	25
2.7. The value of a pilot phase of a trial.....	26
2.8. The need for implementation and process, and cost evaluation	26
3. Design of a potential RCT	28
3.1. Study rationale and background.....	28
3.2. Intervention	28
3.3. Impact evaluation design	28
3.4. Implementation, process and cost evaluation.....	30
3.5. Ethics and data protection	31
3.6. Risks	31
3.7. Timeline	32
4. Conclusion and next steps.....	33
4.1. Key conclusions about feasibility	33
4.2. Key conclusions about design	33
4.3. Next steps for the EUAA and contractor.....	33
ANNEXES	
ANNEX 1: Data collection activities for the Feasibility and Design Study.....	35
ANNEX 2: Data fields needed from the QAT, for all interviews included in the trial	37
ANNEX 3: Indicative data fields needed from relevant data sources relating to applicant/case characteristics.....	39
ANNEX 4: Feasibility checklist.....	40

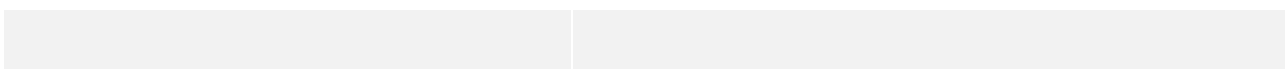
TABLE OF ACRONYMS/ABBREVIATIONS

Acronym/ Abbreviation	Meaning
CEAS	Common European Asylum System
EUAA	European Union Agency for Asylum
EASO	European Asylum Support Office
MESI	Minimum effect size of interest
MDES	Minimum detectable effect size
QAT	Quality Assurance Tool
QED	Quasi-experimental design
RCT	Randomised control trial
ToR	Terms of reference

GLOSSARY

Term	Definition
Cost evaluation	provides essential information about the cost of delivering an intervention – for example, in terms of staff time, materials, and other resources. In the context of this report, it refers to the cost of an average remote interview and an average face-to-face interview.
Fidelity	the extent to which the key components and principles of an intervention are implemented and delivered, in practice, as intended. In the context of this report, it refers to the degree to which remote and face-to-face interviews can be delivered as intended or prescribed.
Impact evaluation	provides insight into outcomes, changes or ‘impacts’ resulting from a treatment or intervention. Impact evaluation aims to establish the cause of observed changes. In the context of this report, it refers to an evaluation which determines the impact of interview modality on interview quality.
Implementation and process evaluation	provides insight into the ‘how and why’ of the results from the impact evaluation which is useful for understanding how the interview process might be improved.
Multi-site trial	a trial in which multiple locations or administrative units are involved, and the trial is run semi-independently across sites.
Power calculations	indicate how large a trial needs to be to detect a particular effect size, known as the minimum effect size of interest (MESI) or indicate the size of effect the trial can detect with reasonable probability, should such an effect exist, known as the minimum detectable effect size (MDES).
Randomisation	the process of randomly allocating study participants to the interventions being tested in the trial. In the case of this Feasibility and Design Study, this would involve random allocation to remote and face-to-face interview modalities.
Randomisation ratio	the proportion of participants who are assigned to one group or another in a RCT.
Randomised control trial	a trial in which participants are randomly assigned to different groups to test a specific treatment or intervention. Differences between the groups are subject to statistical analysis. Abbreviated as RCT.
Stratification	dividing the sample into groups (strata) before randomisation.
Study arm	one of the two or more conditions to which people would be randomised – e.g., the treatment or the control condition.
Vulnerable applicant	in context of this report, this refers to minors whether accompanied or unaccompanied, immediate relatives of victims of sea wrecks (parents, siblings, children and spouses), persons with disability, elderly, pregnant

women, single parents, families with minor children, victims of trafficking, persons with serious illnesses, persons with mental or psychological disabilities and persons who have been subjected to torture, rape or other serious forms of psychological, physical or sexual violence, such as victims of genital mutilation.



EXECUTIVE SUMMARY

Background and Study aims

This report outlines the findings of a Feasibility and Design Study examining if it would be possible to implement a randomised control trial (RCT) based evaluation to compare remote and face-to-face personal interviews in the asylum process. The aim of such a RCT would be to conduct an evaluation of the impact of these different interview modalities on the quality of the interview.

The personal interview is an essential stage in the examination of applications for international protection within the EU. During the interview, details of an applicant's case are explored and information is gathered to allow the relevant national authority to make a decision about whether to grant international protection. Over the years, and in particular since the start of the COVID-19 pandemic, the European Union Agency for Asylum (EUAA or the Agency) and national authorities have started to use remote approaches, conducting interviews through the use of video conference software. The EUAA, in collaboration with partners, has developed practical recommendations on conducting personal interviews remotely. This guidance includes technical, security, confidentiality, and practical arrangements prior to, during and after the interview. The guidance outlines legal considerations, interview techniques as well as quality-related and vulnerability-related considerations.

Even though pandemic-related restrictions are no longer required, countries may consider continuing to use remote interviewing modalities – to complement or replace face-to-face interviews – if interviewing via video link is an effective and efficient *modus operandi* in asylum processes.

As a first step towards building an evidence base about the efficiency and effectiveness of using remote interviewing modalities, the EUAA commissioned this Feasibility and Design Study to:

- Assess the feasibility of running a RCT within the EUAA's operational context and identify the most appropriate treatment and control groups, taking into account relevant features of the two interview approaches.
- Formulate a robust and statistically and ethically sound RCT design framework including an implementation plan.
- Contribute to internal capacity-building activities in the field of RCTs for relevant EUAA staff.

Study methodology

The findings in this report are based on:

- 16 interviews/meetings with 25 interviewees from the EUAA.
- Reviewed key EUAA procedural and guidance documents relating to remote and face-to-face personal interviews, and reports about interview quality.
- Reviewed statistics provided by the EUAA about the number of interviews scheduled/conducted in Cyprus between 2019 and 2022, Greece in 2020 and Malta between 2019 and 2020.¹

The findings in this report are intended to be applicable to any Member State that is considering implementing a RCT but are based on case studies of the processes in Cyprus, Greece and Malta. These countries were selected as, at the time of writing, they were Member States in which the EUAA was providing support through operational plans.² This meant that the EUAA could facilitate access for the Study Team to data and stakeholders from those countries.

Key findings: feasibility

The table below shows the criteria used to assess feasibility and the overall assessment made on each criterion. Overall, the Study finds that it is, in principle, feasible to implement a RCT to compare the

¹ These were the three Member States shortlisted as the sites for a potential RCT during the Study.

² At the time of writing, the Agency had operational plans with Cyprus, Greece, Italy, and Malta for 2022-24, as well as Austria, Belgium, Bulgaria, Czechia, Lithuania, the Netherlands, Romania, Slovenia and Spain covering 2022-23. Latvia had a plan for 2022. A small deployment to Moldova has been carried out in response to displacement from Ukraine, in the first ever deployment of personnel to a non-EU country (according to EUAA website: <https://euaa.europa.eu/news-events/eu-asylum-agency-deploys-moldova>). Cyprus, Greece, and Malta were selected – in discussion with the EUAA – because all three countries had (at the time of the study) EUAA case officers undertaking interviews for eligibility for national protection (in other locations the EUAA supported other kinds of interviews), in relatively large numbers and all three had *some* experience of conducting remote interviews.

effectiveness of remote and face-to-face interviews. A final assessment of feasibility will need to be specific to a Member State in which a trial was to be implemented, taking into account the specific resources and processes in that Member State. This report provides guidance about how a Member State-specific assessment could be conducted.

The feasibility of implementing a trial strongly depends on whether remote interviews have previously been used. If a Member State has not used remote interviewing (or not extensively), the implementation of a trial may require special investment, for example, in installing fast internet connections, procuring, and supplying computers, web cameras, microphones, interview room furniture as well as training case officers to undertake interviews and other requirements.

It is also highly relevant to consider the ethical, social, and political acceptability of a trial in the asylum process, given the vulnerability of some applicants and that this is an issue of political sensitivity. A robust and careful ethics plan, and ethical approval by a recognised ethics body, would be important. The legal framework of the Member State in which a trial is to be implemented should also be examined, to ensure that the trial is compliant with relevant legal standards.

The views of stakeholders would also need to be considered when planning the implementation of a trial – the media, lawyers, non-governmental organisations, applicants, interpreters and officials (among others) all have an interest in such a trial, and their perceptions of the trial are very relevant to successful implementation and to the reputation of the EUAA and national authorities.

Table S1: Summary of feasibility assessment findings

Criteria	Rating of feasibility	Summary of key findings
Is it feasible to define the 'intervention' and implement it with fidelity in operational contexts?	High	Remote and face-to-face interviews have already been used in practice in some countries, showing that it is likely they can be implemented with sufficient fidelity to be tested in a RCT.
Are there feasible options for randomisation?	Medium	In principle it is feasible to randomise at: (i) case officer level (case officers are randomly assigned to conduct either face-to-face or remote interviews for the duration of the trial) or (ii) the applicant level. However, the feasibility of randomisation will depend on the configuration, location, and skills of case officers in the Member State in which a trial would be implemented.
Is it feasible to define a clear primary and secondary research question and to collect data/to answer those questions?	Medium	The Study has identified a primary research question which could feasibly be tested in a potential RCT: 'Are remote and face-to-face interviews equivalent in quality on average?'. The EUAA Quality Assurance Tool (QAT) (or similar tools used nationally) appears to constitute a suitable approach to measuring interview quality in a RCT. However, the following would need to be addressed (for example through a pilot phase of a trial) before the QAT could be used in a trial; the Tool has not been used for the purposes of evaluation (and its statistical robustness has not been tested; modifications would be needed to allow data to be outputted from the Tool in a format that could be used by an evaluation; the use of the Tool depends on the availability of appropriately skilled quality assessors).
Is the number of interviews conducted likely to generate a sufficient sample size within a reasonable timeframe?	Medium	The information needed to make accurate power calculations is not available to the Study Team. This report indicates the information needed to undertake power calculations and determine the sample size for the trial. Internal analysis of the statistical properties of the QAT by the EUAA would likely generate the information needed to undertake indicative sample size calculations specific to a Member State. A pilot phase of a trial would allow sample size calculations to be further refined.
Is it feasible to implement a RCT ethically?	Medium	The Study finds it is feasible for a RCT to be implemented in compliance with legal, ethical, and data protection standards. The Study recommends that vulnerable applicants should be excluded from the RCT.

Is it feasible to implement a RCT in operational, procedural and political contexts?	Medium	It is feasible to implement a RCT, but this depends on Member State-specific factors. The Study sets out criteria to guide the assessment of feasibility in particular Member State contexts.
--	--------	---

Source: Study Team

Key findings: design

Based on the feasibility assessment, this Study outlines a potential design for a RCT which could be feasibly implemented and lead to robust results. Key features of the design are as follows:

- **The focus of the trial:** personal interviews to determine eligibility for international protection, where remote interviews are defined as those in which the interviewing case officer and applicant are not in the same room, and face-to-face interviews are defined as those in which the interviewing case officer and applicant are in the same room.
- The main **research question for the trial** would test the non-inferiority and thus implied equivalence of remote and face-to-face interviewing: 'are remote and face-to-face interviews equivalent in quality on average?'. A suggested secondary question is: 'are any benefits or costs associated with remote personal interviews compared to face-to-face interviews?'
- The trial could include a **process and implementation evaluation** alongside the impact evaluation, in order to develop a better understanding of the experience of stakeholders involved in interviews and explore any positive and negative unintended consequences.
- A **cost evaluation** – to estimate the cost of a typical remote and face-to-face interview – would be an essential element of the trial. If the quality was equivalent across modalities, cost could be a deciding factor on use of remote or face-to-face interviews.
- At least in the first trials, applicants meeting the definition of **vulnerability** in the Member State in question should be excluded from the trial.
- **The approach to randomisation** could involve randomising case officers to either conduct only remote or only face-to-face interviews during the trial and also randomising applicants to an interview modality. This would therefore be called a cluster RCT, where case officers form the clusters and applicants form the cases.
- **The trial would require data** from: the QAT; EUAA (or similar national) case management systems holding information about applicant and case characteristics; national/Agency information about the experience of case officers involved in the trial (or relevant information collected and shared by EUAA personnel under the framework of operational data collections).

Study limitations

A limitation of this Study is that, because historic data about interview quality scores are not available, **the Study Team could not assess the statistical validity of the QAT, nor was it possible to undertake accurate power calculations or sample size estimates.**

Due to this lack of data and given the complexity of implementing a trial in operational contexts, the **Study Team recommends that a pilot phase is undertaken at the start of a trial.** A pilot would provide an opportunity to identify 'real-world' logistical challenges which may not have been possible to anticipate in the design stage, allow ethical risks and issues arising from the assignation of interviews to remote or face-to-face modalities to be identified and be acted on quickly, provide more information on the necessary sample size, and allow an investigation of if (and how) interview quality varies between and within case officers and applicants.

In addition, **the Study Team recommends that an analysis of the statistical properties of the QAT is conducted before piloting a trial.** Without a statistically robust outcome measure, a pilot would not be a good investment of time.

Next steps

If the EUAA, in partnership with a Member State, wished to progress to the next stages of planning a RCT comparing remote and face-to-face interviews, the steps would include:

- Undertaking an analysis of the statistical properties of the QAT (either within the Agency, or by appointing an external contractor).

- Ensuring that the Member State authority understood the full implications of implementing a trial was willing to invest the necessary time and resources and was fully supportive of implementation.
- Appointing a contractor to design and conduct the trial.
- Undertaking a co-design process, led by the contractor, which would assess Member State-specific feasibility and propose a bespoke design for the Member State in question.
- Reviewing and helping to understand data flows and making detailed plans for how the relevant data (including personal, sensitive, and special category data) could be sourced and shared to inform the trial.

1. INTRODUCTION

This is the Final Report for Review produced as part of the Study on the *Feasibility and design for the preparation of a randomised control trial-based evaluation: Remote versus face-to-face asylum personal interviews (Request for services No 46 EUAA/2022/029)* (“The Study”). The Study was conducted by RAND Europe (the Study Team) for the European Union Agency for Asylum (EUAA or the Agency).

1.1. Background to the personal interview and this study

The EUAA was founded in 2010 as EASO (European Asylum Support Office) to support the implementation of the Common European Asylum System (CEAS)³ and promote cooperation and cohesion between Member States.⁴ The Agency began operations under a new name and with an enhanced mandate in January 2022,⁵ in light of the New Pact on Migration and Asylum.⁶ The adjustments to the EUAA’s overarching responsibilities since 2022 include a reinforced mandate to provide assistance to Member States in situations of disproportionate pressure, further promoting EU law and operational standards, and strengthening practical cooperation and information-exchange between Member States.

The personal interview plays a key role in the decision-making process of asylum cases. It is conducted by the personnel of the determining authority, or deployed personnel of the EUAA, with the aim to examine the application for international protection.⁷ Personal interviews are highly sensitive and case officers conducting these interviews need to be trained to allow applicants to present the grounds for their applications in a comprehensive manner⁸ and to ensure that asylum seekers with specific vulnerabilities are able to take part in the interviews without the potential for re-traumatisation and the risk of further harm.⁹ In addition to the case officer and the asylum seeker, an interpreter will be present in the personal interview.¹⁰

COVID-19-related restrictions led to an increase, in some EU countries, in the use of remote interviewing.¹¹ Prior to the pandemic, remote interviews were used in extraordinary cases in some Member States. For some countries, the pandemic was the first time that video technology had been used to conduct interviews. For others, this resulted in an increase in the numbers of cases using video interviews. In 2020,

³ The CEAS established requirements for alignment on migration and asylum policy between EU Member States. Commission of the European Communities. (2008). Communication [...] policy plan on asylum: an integrated approach to protection across the EU. The standards of the CEAS are defined in the Asylum Procedures Directive (2013/32/EU).

⁴ Regulation (EU) No 439/2010 of the European Parliament and of the Council of 19 May 2010 establishing a European Asylum Support Office.

⁵ European Commission. (2022). New pact on migration and asylum: new mandate, new start for the European Union Agency for Asylum. Available [here](#); The 2010 Regulation was repealed by Regulation (EU) 2021/2303 of the European Parliament and of the Council of 15 December 2021 on the European Union Agency for Asylum and repealing Regulation (EU) No 439/2010. L 468/2. which replaced the EASO with the EUAA transforming the EASO into a “fully-fledged agency”.

⁶ EASO. (2021). Asylum report: annual report on the situation of asylum in the European Union. p.5.

⁷ Directive 2013/33/EU of the European Parliament and of the Council of 26 June 2013 on common procedures for granting and withdrawing international protection (recast) Article 14.

⁸ Directive 2013/32/EU of the European Parliament and of the Council of 26 June 2013 on common procedures for granting and withdrawing international protection (recast) Article 15.

⁹ Specific vulnerabilities may include their age, gender, sexual orientation, gender identity, disability, serious illness, mental disorders or as a consequence of torture, rape or other serious forms of psychological, physical or sexual violence, as set out in Directive 2013/33/EU of the European Parliament and of the Council of 26 June 2013 on common procedures for granting and withdrawing international protection (recast) Recital 29.

¹⁰ Access to interpretation is a guarantee that needs to be safeguarded according to Article 12 of the Directive 2013/33/EU of the European Parliament and of the Council of 26 June 2013 on common procedures for granting and withdrawing international protection (recast) Article 12 and Article 15. The interpreter’s role is to facilitate communication in the preferred language of the applicant and the applicant has the right to request the gender of the interpreter.

¹¹ EUAA. (2022). Invitation to Tender: Feasibility and Design for the preparation of a randomised control trial-based evaluation: Remote versus face-to-face asylum personal interviews.

the EUAA issued guidance on how to maintain the standards of personal interviews whether they are conducted in person or remotely.¹²

There is limited evidence about the extent to which remote interviews are effective. There are plausible hypotheses for and against the assumption that remote interviews can be as good as face-to-face interviews in terms of interview quality. Internal EUAA data indicate that the following could be challenging in remote interviews:

- Building a positive relationship with the applicants, particularly vulnerable ones.
- Offering response to medical or safety incidents.
- Connection problems.
- Identifying certain vulnerability indicators: serious illness, mental disorders, victims of violence, single parents, and pregnancy.

On the other hand, internal EUAA data indicate that remote interviews were conducted as planned, and that remote settings:

- Allowed for the creation of a positive atmosphere.
- Did not affect the ability to explore credibility.
- Allowed interviewers to follow up on the information shared by the applicant and maintain control of the interview.
- Allowed applicants and interpreters to participate in the interview from a suitable environment and were provided with appropriate equipment to guarantee confidentiality standards were met.
- Were equally good in comparison to face-to-face in terms of the quality of the interpretation.
- Were equally good in comparison to face-to-face in relation to identifying vulnerable applicants.

These insights have limitations, as they are drawn from EUAA experience with one Member State at one time point. However, they are relevant to the Feasibility and Design Study because they provide a plausible reason to believe that, in the right conditions, remote interviews can deliver similar outputs and outcomes as face-to-face interviews. This conclusion can also be drawn from the information contained in guidelines and recommendations on the implementation of remote interviews developed by the EUAA and United Nations High Commissioner for Refugees (UNHCR) (which, by definition, indicate that remote interviews can be as good as face-to-face interviews).¹³

Given there is no firm evidence that remote compared to face-to-face interviews are 'better', the EUAA commissioned this Feasibility and Design Study to investigate whether it would be possible to design a robust evaluation, employing a randomised design, to test the equivalence of remote versus face-to-face personal interviews as alternative approaches to meeting minimum requirements of the CEAS.¹⁴

1.2. Objectives of the Feasibility and Design Study

This Feasibility and Design Study has three primary objectives, as defined in the Terms of Reference (ToR):

- **Objective 1:** Contribute to evidence of promise on the equivalence of remote versus face-to-face asylum determination personal interviews as alternative approaches to meeting minimum CEAS¹⁵ requirements.
- **Objective 2:** Further enhance evidence-based knowledge within the EUAA in relation to face-to-face versus remote asylum personal interviewing.

¹² EASO. (2020). EASO Practical recommendations on conducting the personal interview remotely.

¹³ European Union Agency for Asylum (2020). Practical recommendations on conducting the personal interview remotely. Available at: <https://euaa.europa.eu/publications/practical-recommendations-conducting-personal-interview-remotely>

UNHCR (2020). Remote Interviewing: Practical Considerations for States in Europe. Available at: <https://data.unhcr.org/en/documents/download/77134>

¹⁴ European Commission. 'Common European Asylum System (CEAS)'. https://hoe-affairs.ec.europa.eu/pages/glossary/common-european-asylum-system-ceas_en

¹⁵ European Commission. 'Common European Asylum System (CEAS)'. https://hoe-affairs.ec.europa.eu/pages/glossary/common-european-asylum-system-ceas_en

- **Objective 3:** Provide on-the-job learning opportunities and a demonstration experience for EUAA staff to undertake RCT evaluations as a means to produce evidence to support fulfilment of the EUAA's mandate.

To achieve these objectives, the ToR outlined the following tasks:

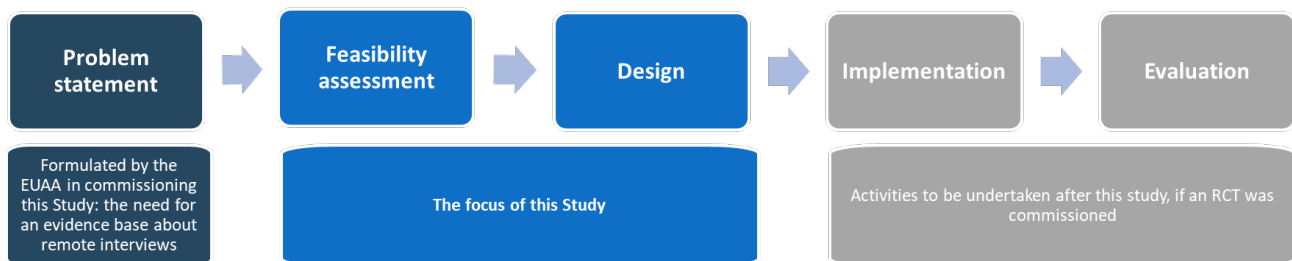
- Assess the feasibility of running a RCT within the EUAA's operational context and identify the most appropriate treatment and control groups, taking into account relevant features of the two interview approaches.
- Formulate a robust and statistically and ethically sound RCT design framework including an implementation plan.
- Contribute to internal capacity-building activities in the field of RCTs for relevant EUAA staff.

The Study commenced in August 2022 and was completed in February 2023.

Figure 1 shows the steps typically involved in developing an evaluation using a randomised design. Step 1 – the problem statement – was formulated by the EUAA in commissioning this study: the EUAA would like to build an evidence base to support decision making about whether there is a case for conducting remote interviews, alongside face-to-face interviews, as a regular practice rather than solely in extenuating circumstances. This Study focuses on steps 2 and 3 of the process (in blue) in Figure 1:

- To assess whether it is feasible to conduct a RCT to assess remote compared to face-to-face asylum personal interviews.
- To propose a potential design for such a RCT.

Figure 1: Steps involved in developing a RCT, and the focus of this study



Source: Study Team

Feasibility and Design Studies are conducted to find out if, how, and at what cost an evaluation using a randomised design could be implemented in an operational context and if a RCT would produce a valid and reliable result. The ambition of this Study was to arrive at:

- A clear description of 'what, who, where, when' of remote and face-to-face interviews
- A clear evaluation question to be addressed in a RCT
- An understanding of the data that would allow the evaluation question to be answered
- An understanding of the sample sizes and power
- A proposed approach to randomisation
- A thorough exploration of data protection and ethical issues
- An estimate of how long a RCT would take.

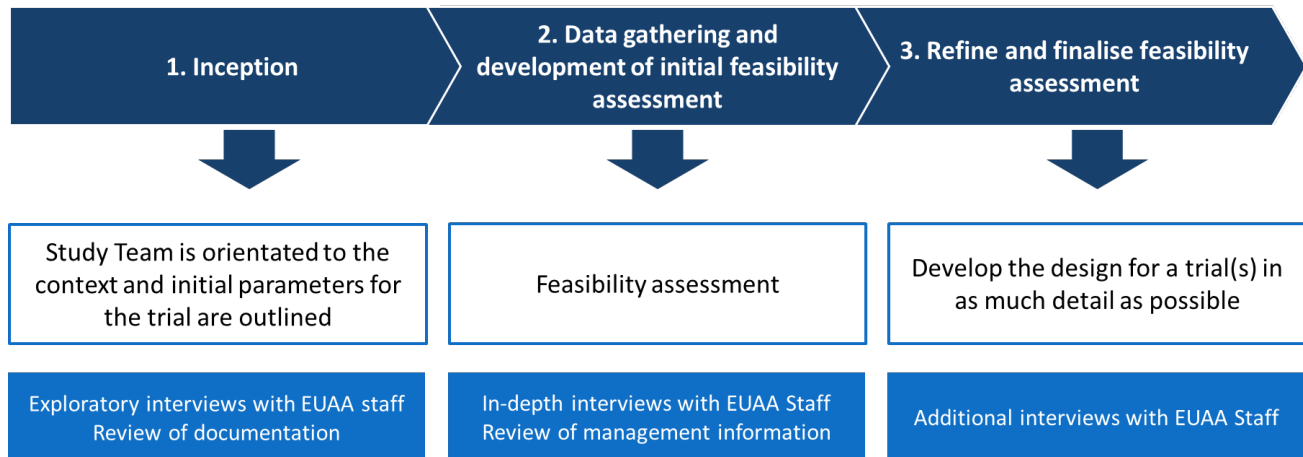
The Study aimed to deliver a realistic assessment of the risks and potential of an evaluation using a randomised design so that the EUAA can make informed decisions about if and how they could use this approach.

To achieve Objectives 2 and 3, the Feasibility and Design Study was used as a 'demonstration project' through which Agency staff could learn about RCT design. During the Study, two capacity-building workshops were delivered, during which the Study Team explained the feasibility and design process and the rationale for design choices.

1.3. Methodological approach and the data collection that has informed this Study

The study was divided into three phases, the aims and outputs of which are outlined in Figure 2.

Figure 2: The approach to the Feasibility and Design Study



Source: Study Team

The findings in this report are intended to be applicable to any Member State considering implementing a RCT but are based on case studies of the processes in Cyprus, Greece, and Malta. These countries were selected – in discussion with the EUAA – because, at the time of the Study (August 2022 to February 2023) they were Member States in which the EUAA provided support through operational plans¹⁶; they were where the EUAA had case officers undertaking interviews for eligibility for national protection (in other locations the EUAA supported other kinds of interviews) in relatively large numbers; and all three had some experience of conducting remote interviews.

Therefore, the EUAA could facilitate access of the Study Team to data and stakeholders about processes in those countries.

The findings in this report are based on data synthesised from interviews with EUAA staff and a review of documentation. Annex 1 provides an overview of the interviews and documentary sources utilised to produce this report. In brief, the Study Team:

- Conducted 16 interviews/meetings with 25 interviewees.
- Reviewed key EUAA procedural and guidance documents relating to remote and face-to-face personal interviews, and reports about interview quality.
- Reviewed statistics provided by the EUAA about the number of interviews scheduled/conducted in Cyprus between 2019 and 2022, Greece in 2020, and Malta between 2019 and 2020.¹⁷

1.4. Study strengths and limitations

This Final Report outlines the findings from the feasibility assessment and proposes a design for a potential RCT comparing remote and face-to-face interviewing, covering all aspects outlined in the ToR.

¹⁶ At the time of writing, the Agency had operational plans with Cyprus, Greece, Italy and Malta for 2022-24, as well as Austria, Belgium, Bulgaria, Czechia, Lithuania, the Netherlands, Romania, Slovenia and Spain covering 2022-23. Latvia had a plan for 2022. A small deployment to Moldova has been carried out in response to displacement from Ukraine, in the first ever deployment of personnel to a non-EU country (according to EUAA website: <https://euaa.europa.eu/news-events/eu-asylum-agency-deploys-moldova>).

¹⁷ These were the three Member States shortlisted as the sites for a potential RCT during the Study.

Because there is limited historic data about the main proposed outcome measure (interview quality scores), the report is not as detailed as planned in relation to sample sizes and power calculations for a potential RCT.

In this report, where it is not possible to specify an aspect of the design without reference to a particular Member State, relevant considerations and implications are outlined.

1.5. The aim and structure of this Final Report

This report describes, in as much detail as possible, the findings from the feasibility assessment and outlines the design of a potential RCT with the rationale for design choices.

Chapter 2 outlines the assessment of feasibility, **Chapter 3** sets out a design for a possible RCT and an implementation, process and cost evaluation to be conducted alongside a RCT, and **Chapter 4** provides a conclusion and outlines next steps.

2. FEASIBILITY OF A RCT

This study has looked at the following six criteria to determine the feasibility of a RCT:

- Is it feasible to define the 'intervention' and implement it with fidelity in operational contexts?
- Are there feasible options for randomisation?
- Is it feasible to define a clear primary and secondary research question and to collect data to answer those questions?
- Is the number of interviews conducted likely to generate a sufficient sample size within a reasonable timeframe?
- Is it feasible to implement a RCT ethically?
- Is it feasible to implement a RCT in social and political contexts?

Findings in relation to each of these criteria are summarised in Table 1 and explained in the sub-sections below. **Overall, the Study finds it is, in principle, feasible to implement a RCT, although final assessment of feasibility would need to be specific to a Member State in which a trial was to be implemented.**

The Study Team recommends that an implementation, process, and cost evaluation is undertaken alongside a trial – this is described in Section 3.4.

Table 1: Summary of Study findings on feasibility

Criteria	Assessment of feasibility	Summary of key findings and recommendations from this Study
Is it feasible to define the 'intervention' and implement it with fidelity in operational contexts?	High	Remote and face-to-face interviews can be defined with sufficient precision to be tested in a RCT. Remote and face-to-face interviews have already been used in practice in some countries, showing that it is likely that they can be implemented with sufficient fidelity to be tested in a RCT.
Are there feasible options for randomisation?	Medium	In principle it is feasible to randomise: (i) at case officer level (case officers are randomly assigned to conduct either face-to-face or remote interviews for the duration of the trial) or (ii) at the applicant level. However, the feasibility of randomisation will depend on the configuration, location, and skills of case officers in the Member State in which a trial would be implemented.
Is it feasible to define a clear primary and secondary research question and to collect data/to answer those questions?	Medium	The Study has identified a primary research question that could feasibly be tested in a potential RCT: 'Are remote and face-to-face interviews equivalent in quality on average?' The EUAA Quality Assurance Tool (QAT) appears to constitute a suitable approach to measuring interview quality in a RCT. This Tool has some advantages: it is already developed and accepted; it is used in practice; and it captures most relevant aspects of interview quality. However, the QAT has not been used for the purposes of evaluation and reporting functions in the EUAA App currently do not allow data from the Tool to be outputted in a format that could be used by an evaluation (although in principle this is not complicated to do). Because data are not available, the statistical robustness of the Tool would also need to be tested. Feasibility would also depend on the availability of appropriately skilled quality assessors. Similar tools used by national authorities might also be used.
Is the number of interviews conducted likely to generate a sufficient sample size within a reasonable timeframe?	Medium	The information needed to make accurate power calculations is not available to the Study Team. Indicative power calculations have been made based on some untested assumptions and should therefore be treated with caution.

Criteria	Assessment of feasibility	Summary of key findings and recommendations from this Study
Is it feasible to implement a RCT ethically?	Medium	While recognising the vulnerability of applicants for asylum, the Study finds it is feasible for a RCT to be implemented in compliance with legal, ethical, and data protection standards. However, an ethical protocol would need to be prepared, specific to the Member State in which a RCT was implemented, and would need to be developed by the contractor appointed to conduct a RCT.
Is it feasible to implement a RCT in operational, procedural and political contexts?	Medium	It is feasible to implement a RCT, but this depends on Member State-specific factors. The Study sets out criteria to guide the assessment of feasibility in particular Member State contexts, including taking into consideration practical, financial, social, and political context.

Source: Study Team

2.1. Is it feasible to define the 'intervention' and implement it with fidelity in EU asylum determination contexts?

For a RCT to be feasible it is essential that both the intervention(s) or practices being tested and the control (or 'business as usual') conditions are thoroughly understood and described. This includes the stakeholders involved and the wider context and processes in which the practice sits. The 'programme theory' of the intervention should also be defined. A programme theory explains how an intervention is understood to contribute to a chain of results that produce the intended or actual impacts.¹⁸

For a RCT to be feasible it is important that an intervention can be delivered in practice with fidelity in the operational context. In the context of RCTs and evaluations the term 'fidelity' refers to the degree to which the intervention is delivered as intended or prescribed.¹⁹

2.1.1. Specifying the types of personal interview included in the RCT

It is recommended that the focus for a RCT is the use of personal interviews to determine **eligibility for international protection**.

Personal interviews may also be conducted to determine the *admissibility* of a claim for international protection (i.e., a step before eligibility is determined). The Study Team recommends that a RCT comparing remote and face-to-face interviews should focus exclusively on personal interviews to inform claims for eligibility because interviews for admissibility are conducted for only certain types of applicants (e.g., those from safer countries of origin list where an answer can be given quickly). Additionally, the recommended outcome measure (see Section 2.3.2) does not apply to admissibility interviews.

2.1.2. Defining 'remote' and 'face-to-face' interviews

The following definitions are proposed by the Study Team:

- Remote interviews are those in which the interviewing case officer and applicant are **not** in the same room.
- Face-to-face interviews are those in which the interviewing case officer and applicant **are** in the same room.

¹⁸ Better Evaluation, n.d., 'Develop programme theory/theory of change', <https://www.betterevaluation.org/frameworks-guides/rainbow-framework/define/develop-programme-theory-theory-change>, Accessed: 30/01/2023.

¹⁹ Berkel, Mauricio, Schoenfelder, & Sandler, 2011. *Putting the Pieces together: An Integrated Model of Program Implementation*. Prev Sci. 2011 Mar;12(1):23-33. doi: 10.1007/s11121-010-0186-1.

The location of interpreters and lawyers is not central to these definitions. However, during the personal interview process it is never acceptable (in any Member State) for an interpreter and an applicant to be left alone in the same room. This is for reasons of protecting applicants, to ensure the interpreter does not influence the dynamics of the interview and to avoid the interpreter being placed in a difficult situation by the applicant (e.g., the applicant asking the interpreter for help). Therefore, in the 'remote' condition, there are two possible configurations:

- Applicant in room 1; case officer and interpreter in room 2.
- Applicant in room 1; case officer in room 2; interpreter in room 3.

In the face-to-face condition, the possible configurations are:

- Applicant and case officer in room 1; interpreter in room 2.
- Applicant, case officer and interpreter in room 1.

Box 1: Terminology

In the context of a potential RCT comparing remote and face-to-face interviewing, the terms 'intervention' and 'business as usual' are potentially misleading since remote interviews are already used in some countries (such as Greece). In other words, remote interviews are 'business as usual', to some extent in some jurisdictions. The terminology of 'modality', 'condition', or 'practice' might be more appropriate.

If implementing a RCT in a Member State where remote interviews were not conducted, it might be appropriate to refer to remote interviews as an 'intervention' – since the implementation of the trial would result in a different practice.

2.1.1. Articulating the 'programme theory' of remote and face-to-face personal interviews

A programme theory explains how an intervention (a project, a programme, a policy or a strategy) is understood to contribute to a chain of results that produce the intended or actual impacts.²⁰ It is good practice in assessing the feasibility of an evaluation to define how the intervention is intended to lead to results. In the context of this Study, a programme theory sets out all the relevant inputs, activities and outputs involved in conducting face-to-face and remote interviews, which are important to interview quality and to the ultimate aim that the interview allows for a correct and effective decision regarding eligibility for international protection.

The value of the programme theory for this Feasibility and Design Study is that it clearly sets out the conditions (inputs and activities) that play a role in the delivery of both interview modalities with fidelity (discussed in the following section), which might be monitored as part of a trial.

Box 2: Programme theory for remote and face-to-face personal interviews

Inputs: Trained and competent case officers; trained and competent interpreters; interview rooms; internet connection; functioning laptop, microphone, camera; connectivity and IT; information about the applicant collected at registration; scheduling and communicating about the interview; a case file containing relevant information; support staff to support applicant on day of the interview; transport to the interview.

Activities: The interview 'conversation(s)'; exchange of relevant documentation.

Output: Interview transcript; updated case file.

Outcome: A quality interview; relevant information is gathered.

Result: The interview allows for effective and correct decision regarding eligibility for international protection.

Facilitators/barriers/causal mechanisms: Effective interviewing/questioning techniques/skills of case officers; quality of interpretation; the development of rapport and trust between interviewer and applicant; provision of supporting documentation; quiet, safe, comfortable surroundings; applicant feels safe; lack of backlogs and delays in processing applications.

²⁰ BetterEvaluation, n.d., 'Develop programme theory/theory of change', <https://www.betterevaluation.org/frameworks-guides/rainbow-framework/define/develop-programme-theory-theory-change>, Accessed: 30/01/2023.

2.1.2. Assessing whether both interview modalities can be implemented with fidelity

In the context of developing RCTs, the term 'fidelity' means the degree to which the intervention is delivered as intended or prescribed. In the case of remote and face-to-face interviews, fidelity means the extent to which both modalities could be implemented in line with EUAA guidance and in accordance with the Asylum Procedures Directive (2013/32/EU).

Both face-to-face and remote interviews have been implemented in practice (in Greece, Malta and Cyprus), suggesting that, with appropriate infrastructure and monitoring, it is likely that the modalities can be implemented with fidelity.

The degree of fidelity is dependent on the particular situation in a given Member State. Relevant features that need to be present for a remote interview to be implemented with fidelity are set out in Table 2, drawing on the programme theory.

The extent to which a RCT can be implemented with fidelity depends on the resources available in each country. **If a Member State has not used remote interviewing extensively, the implementation of a RCT may require special investment to ensure these features are present.** For example, installing fast internet connections, procuring and supplying computers, web cams, microphones, interview room furniture etc. The guidelines produced by the EUAA provide information about the equipment and practicalities that need to be in place to implement a remote interview with fidelity.²¹

In selecting a Member State in which to implement a trial, the factors outlined in Table 2 should be considered. The extent to which these barriers to implementation of remote and face-to-face exist could be relevant to whether a Member State is a suitable site for a trial and is relevant to understanding which additional investment might be needed to facilitate implementation.

Table 2: Factors relevant to fidelity of delivery

Factor relevant to fidelity of delivery	How it is relevant to fidelity of delivery
Enough support staff	A member of staff to: greet the applicant, escort them to the interview room, set up the interview connection (i.e., over Microsoft Teams/Zoom), address technical difficulties, provide assistance if the applicant becomes unwell or distressed, and collect and share documents produced by the applicant in support of their case. ²²
Quality of IT and connectivity	A consistent internet/Wi-Fi connection. A clear, good quality video and audio connection between the applicant, case officer, and interpreter.
Enough interview rooms	An appropriate number of interview rooms with relevant IT equipment. For example, in Cyprus and Malta, at the time of writing, facilities were not available for remote interviews. To implement a RCT in such contexts, additional interview rooms and equipment would be required. In contrast, there were at least 10 or 11 rooms in which remote interviews can be conducted in Greece.
Transport	Transport to and from the interview location for applicants, case officers, and interpreters.
Digitalisation of record keeping/ case management	The extent to which the interview and decision processes in the Member State rely on paper records and case files, or digital files. For example, in Cyprus, there is significant use of paper-based systems, including 'wet' signature needed from case officers after interviews are completed. This could add time and cost implications to implementing remote interviews.
Enough trained case officers	Case officers with appropriate training to undertake remote and face-to-face interviews.

²¹ European Union Agency for Asylum (2020). Practical recommendations on conducting the personal interview remotely.

²² The Study Team understands that flow managers could assist in this capacity.

	<p>Enough case officers. This is relevant because it will affect the number of personal interviews that can be conducted, per week or month, in a RCT, and therefore how long the RCT will need to operate to reach a sufficient sample size.</p> <p>By nature, asylum services are reactive and adaptive, and it might not always be known in advance how many case officers will be in post in the future. In estimating future numbers of case officers, it should be considered that EUAA case officers (in the situation where EUAA officers provide operational support) might be on short-term contracts and/or have short contractual notice periods.</p> <p>For instance, in Malta, some case officers are on interim contracts with a two-week notice period and can request leave with one-week's notice.</p> <p>Case officer turnover during the implementation of a trial could pose risks to the fidelity of implementation, or cause delays in the timeline of a trial, if understaffing meant fewer interviews were conducted, taking longer to reach the required number of interviews for a robust analysis.</p> <p>In some countries it might be necessary to consider employing additional case officers for the duration of the trial.</p>
Backlog and system pressures, understaffing	<p>If a national asylum agency is under pressure to clear case backlogs, or is experiencing staff shortages, this may not be a suitable environment in which to implement a RCT – which would require time and resources, the implementation of new processes etc.</p>

Source: Study Team drawing on interview findings

2.1.3. Understanding the role of the stakeholders involved and how they might affect feasibility

As part of a feasibility study, the description of the practice being tested in a potential RCT should include a description of the organisations and individuals involved in its delivery. Regardless of the Member State in which a trial is conducted, similar stakeholders are likely to be involved in the personal interviews to determine eligibility for international protection. They are listed below.

When considering implementing a trial, the EUAA, in partnership with the relevant national authority, will need to consider what and how to communicate to these stakeholder groups about a trial. These stakeholder groups might have concerns about a trial – even one which was carefully planned and has ethical approval.

2.1.4. Case officers

Personal interviews are conducted by case officers (sometimes called case workers – the name varies by Member State), who are responsible for interviewing the applicant, reviewing the application and supporting documents, and drafting a written recommendation on whether the asylum application should be accepted.

The EUAA provides support to some Member States by providing case officers who undertake personal interviews.²³ Case officers from national authorities carry out similar tasks to EUAA case officers, except that national case officers can draft an official (first instance) decision on the case, rather than make a recommendation about the decision (which is the role of the EUAA case officers).

Conducting personal interviews is a highly skilled task. EUAA staff were of the view that experience was important, with case officers with the most experience being the most proficient interviewers. Because case officer experience is a variable likely to affect the primary outcome of a potential RCT (the quality of a personal interview), this is a possible stratification variable (see Section 2.2.3) and a trial would need to monitor the experience of case officers involved. An assessment of case officer experience could be made via calculating years of experience, extent of training undertaken and/or total case load over time.

²³ The direction and control of EUAA case officers is different in Greece compared to Cyprus and Malta. In Greece, an 'embedded model' involves the national authorities being responsible for the day-to-day tasks of case officers, with the EUAA supporting the activities in accordance with guidelines from the central level. In the embedded model, case officers are seconded by the EUAA to national authorities of the host Member State and perform their work on the premises of the relevant national authorities. In comparison, case officers in Cyprus and Malta are under the direct control of the EUAA. However, although the EUAA has less control over case officers compared to non-embedded deployments, EUAA staff interviewed for this report were of the view that the EUAA has enough involvement in the employment and day-to-day work of EUAA case officers, even where they are embedded in a Member State asylum service, to control the implementation of a RCT.

The EUAA and relevant national authorities may have to make a choice about whether a future RCT involves both EUAA case officers and case officers from the national authority. The factors to consider in making this decision are outlined in Box 3.

Box 3: Should an evaluation involve case officers from the EUAA and/or the national authorities?

Reasons to include both EUAA and national authorities case officers	Reasons to include only EUAA case officers	Reasons to include only national authorities case officers
<p>The results of a RCT would be more relevant/ applicable to both the EUAA and Member States.</p> <p>Including both would likely mean that more case officers could be involved in the RCT, which would lead to a larger sample size and/or a quicker trial.</p>	<p>If the EUAA is implementing the trial, it might be easier to direct and monitor EUAA case officers' activities and specify how they should conduct themselves in the trial. In turn, this makes it more likely that the interview modalities would be delivered with fidelity.</p> <p>An additional benefit of only involving EUAA case officers in the trial is that there would be easier access to necessary statistical data, through the operational data collected by the EUAA.</p> <p>However, the extent to which the EUAA manages their case officers depends on the operational model.</p>	<p>Using only EUAA case officers in a RCT, and not those of Member States, may have implications for the interpretation of results since these might not be applicable/generalisable to Member State case officers. Depending on the extent to which the practices of Member State case officers differ from those of the EUAA officers, interviews conducted by a national case officer could be a different 'intervention' which should be evaluated separately.</p> <p>National authorities might find the results more relevant if the trial included only national officers.</p> <p>EUAA case officers may conduct, on average, fewer interviews per week compared to case officers from national authorities. This could increase the time needed to deliver the trial.</p>

Source: Study Team

2.1.5. Interpreters

In most cases, an interpreter will be present in the personal interview. The interpreter's role is to facilitate communication in the preferred language of the applicant.²⁴

Interpretation services are procured by national authorities or the EUAA via contracted agencies. There is a framework contract on interpretation and security checks are carried out.

The quality of interpretation is a key factor determining the quality of a personal interview. The level of training and/or experience among interpreters varies and can be challenging to monitor and measure. Since interpretation is a variable that can affect the primary outcome of a potential RCT (the quality of a personal interview), consideration should be given to if and how interpretation quality can be measured during a trial.

2.1.6. Schedulers

The process for scheduling interviews differs between Member States, but this role is crucial for the design of the trial as they would be likely to be responsible for randomisation (discussed further in Section 2.2).

The implementation of a RCT would involve **inserting new steps into the process of scheduling personal interviews**. Steps including, for example, assessing whether an applicant met any inclusion/exclusion criteria for involvement in the trial; undertaking randomisation to assign the applicant to either remote or face-to-face modalities. Box 4 illustrates the similarities and differences in scheduling processes in Cyprus, Greece, and Malta.

²⁴ Pursuant to articles 12 and 15 of Directive 2013/33/EU of the European Parliament and of the Council of 26 June 2013 on common procedures for granting and withdrawing international protection (recast).

Box 4: Processes for scheduling interviews²⁵: cases of Cyprus, Greece, and Malta

Cyprus EUAA personnel schedule interviews. Interview scheduling is based on numbers of interpreters, rooms available and monthly team composition. The national authority provides the EUAA with case files and the EUAA sorts the applicants based on nationality and allocates cases to case officers. Scheduling is different in Pournara and Nicosia. For Nicosia there are more files to plan further ahead whereas in Pournara the number of cases is not known in advance as much.	Greece National authorities schedule interviews. At the registration stage, applicants are notified in what regional asylum office their interview will be conducted. Interview scheduling is based on citizenship (as there are some specific locations/asylum units conducting interviews only for specific citizenships) and on the availability of interpreters of the specific languages of the applicants. At the lodging of applications, applicants are provided with an interview date, they are informed where their interview will take place and whether it is remote or face-to-face.	Malta EUAA personnel schedule interviews. Interviews are organised two weeks in advance. There are three flow managers dedicated to scheduling. Flow Managers have information about the number of case officers available, they know what interpretation resource will be available and are responsible for selecting applicants from the backlog, to fit them with availability of case officers. Flow managers receive instructions from the Malta Asylum Service on which applicants to prioritise and how to address any backlog.
--	---	---

Source: Study Team drawing on interview findings

2.1.7. Support staff

For remote interviews, at least one member of staff will be required to greet the applicant, escort them to the interview room, set up the interview connection (i.e., over Microsoft Teams/Zoom), be available to address technical difficulties and in the event that the applicant becomes unwell or distressed, and to collect and share documents produced by the applicant in support of their case.²⁶

2.1.8. Lawyers

A lawyer may be present. In Greece, the Study Team understands that sometimes applicants have a lawyer provided by a non-governmental organisation – who often take up cases of vulnerable people (it is rare for applicants to hire a private lawyer). However, in Cyprus and Malta it is not common for a lawyer to attend the personal interview. In Malta, national rules allow a lawyer to observe but not to intervene.

EUAA interviewees reported that lawyers have not previously objected to EUAA staff observing or ‘shadowing’ interviews, and that lawyers generally welcome remote interviews if these shorten the time between registration and an interview.

2.1.9. Applicants

Applications can be made by individuals, groups and families. In the case of an application where individuals are linked (e.g., a family), it is usual practice for adults and minors over 15 to each be interviewed separately, but not minors under 15.

Some applicants are classified as ‘vulnerable’. The definition of vulnerable applicants in the Asylum Procedures Directive is as follows: “minors, unaccompanied minors, disabled people, elderly people, pregnant women, single parents with minor children, victims of human trafficking, persons with serious illnesses, persons with mental disorders and persons who have been subjected to torture, rape or other serious forms of psychological, physical or sexual violence, such as victims of female genital mutilation”.²⁷

²⁵ The below refer to interviews conducted by EUAA case officers only.

²⁶ The Study Team understands that flow managers could assist in this capacity.

²⁷ Directive 2013/33/EU of the European Parliament and of the Council, Chapter IV Provisions for Vulnerable Persons. <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013L0033&from=EN>

Box 5 gives an example of how vulnerable applicants are identified in Greece. In Greece, vulnerable applicants are not eligible for remote interviews.

Box 5: Definition and process for identifying vulnerable applicants in Greece

The definition in Greek legislation of a vulnerable applicant is: “minors, accompanied or unaccompanied, immediate relatives of victims of sea wrecks (parents, siblings children and spouses), persons with disability, elderly, pregnant women, single parents, families with minor children, victims of trafficking, persons with serious illnesses, persons with mental or psychological disabilities and persons who have been subjected to torture, rape or other serious forms psychological, physical or sexual violence, such as victims of genital mutilation”.²⁸

The determination of the status of vulnerable persons is carried out after a personalised assessment of each case.

Reception or Asylum Authorities can refer an applicant to public hospital doctors, hospitals or public mental health facilities, specially contracted doctors, the Medical Control and Psychosocial Support of Reception and Identification Centers or to the Closed Controlled Structures, to assess vulnerability. This assessment is used for the immediate coverage of special reception needs, as well as for the provision of special procedural guarantees in the context of their asylum procedure.

Greece has procedural guarantees that mean that vulnerable applicants, where the EUAA or the Asylum Authorities has prior information about their vulnerability, do not have remote interviews.

According to standard operating procedures, vulnerable applicants (provided that the vulnerability is determined/known in advance e.g., at stage of registration) have their interviews conducted face-to-face and are not assigned to remote interview.

However, vulnerability is not always detected in advance, and might come to light during an interview. If the interview is being conducted remotely, the case officer uses professional judgement to decide whether to stop the remote interview, and re-arrange for a face-to-face interview, or to continue remotely.

Source: Study Team drawing on interview findings

This Study recommends that applicants identified as vulnerable are not included in a RCT. There is no robust evidence about how vulnerable applicants experience remote compared to face-to-face modalities. It is not known if remote interviews are more difficult or stressful for vulnerable applicants. It is plausible that some vulnerable applicants may find the ‘distance’ of a remote connection a more comfortable situation in which to disclose information.

However, there is a strong perception among the EUAA stakeholders consulted for this Study that remote interviews are not suitable for vulnerable applicants, and that assigning vulnerable applicants to remote interviews raises ethical concerns. If a trial did include vulnerable applicants, this could act as a barrier to securing the full support of national agencies and others – which is essential.

Another reason for not including vulnerable applicants is that the trial aims to test these interview modalities *as they are currently used* in the specific Member State where an evaluation was being implemented. In a Member State such as Greece, vulnerable applicants are not assigned to remote modalities.

The trial protocol should specify what should happen in a case where evidence that an applicant is vulnerable is not detected in advance, and only comes to light during an interview, including the conditions in which the interview should be stopped (if it is using a remote modality) and whether the applicant should be excluded from the trial.

Some cases are considered to be complex. The factors that might lead an applicant’s case to be classified as complex include where: a case involves serious crime; the authority is not satisfied with the decision after the first interview; or the applicant is not fit for interview (mentally, medically, etc.), and there is a need to reschedule the interview. Complex cases can also result from applicants being from specific countries of origin, or from particular regions in a given country, for which national practices assign longer interview slots. The need for or ability of applicants to gather additional information proceeding the interview may also add complexity to the case. Ultimately, the categorisation of an application as complex will be based on a case-by-case judgement of a case officer and other staff, given the available information about an applicant’s circumstances.

This study recommends that complex cases are included in a trial, because such cases are currently eligible for remote interviews (e.g., in Greece). During interviews, EUAA staff were of the view that complex cases could be included in a trial. In theory, random assignment should result in similar numbers of complex cases in both ‘arms’ of the trial (i.e., one arm of the trial is the remote modality, the other is the face-to-face modality).

²⁸ Definition provided by EUAA staff in Greece.

2.1.10. Line managers and quality reviewers

Case officers – from the EUAA or national authorities – have supervisors and line managers who oversee their work. As discussed further in Section 2.3.5, quality reviewers may have an important role in a trial, since they could be responsible for assessing the quality of interviews.

2.2. Are there feasible options for randomisation?

Randomisation refers to the process of randomly allocating study participants to the interventions being tested in the trial. In the case of this Feasibility and Design Study, this would involve random allocation to remote and face-to-face interview modalities.

Randomisation is a key part of carrying out a RCT and allows the researcher to test if the intervention causes a different outcome to the control.

Randomly assigning those in the trial to either face-to-face or remote interviews will minimise the influence that observable and unobservable factors have on the quality of the interview.

The Study has focused on two short-listed possible approaches to randomisation:

- Randomisation at the **level of the applicant only** – where individual applicants are assigned to either the intervention (remote interview) or 'business as usual' (face-to-face interview).
- Randomisation at the **level of the case officer** facilitating the interview – case officers are assigned to either the intervention (remote interview) or business-as-usual (face-to-face interview) for the duration of the trial, *and then applicants are also* randomly assigned to case officers.

The Study Team recommends the second option: a cluster RCT with randomisation at the level of the case officer to form the clusters, as well as randomisation at the level of the applicant. The advantages of this approach are:

- It would reduce the risk of contamination – the risk that the case officers' experience of conducting face-to-face interviews during the trial affects the ways in which they conduct remote interviews (or vice versa).
- It might more closely resemble 'real life practice' as case officers would become practiced at implementing remote interviews.
- Participants would also be randomised to case officers to account for possible variations in caseload characteristics (e.g., case complexity, applicant characteristics, priority, etc).

2.2.1. Feasibility of randomising case officers

The Study Team concludes that randomising case officers is, in principle, feasible, but the approach to randomisation would need to be defined in relation to the specific Member State arrangements relating to, for example, the location where case officers are based and whether there are a sufficient number of case officers with the relevant skills and experience to undertake remote interviews. It would also depend upon support for randomising case officers from the national authorities.

In the case of Greece, there have been periods of time in which some case officers have been assigned to conduct only remote or only face-to-face interviews, which indicates that randomising at the case officer level would be possible. The randomisation process could be undertaken using an app, website, or spreadsheet that generates the random allocation. Such tools are commonly used in RCTs.

The Study Team recommends that randomisation of case officers is undertaken once, at the start of the trial. It is likely that some case officers might leave their posts during the trial, as some staff turnover is to be expected. To account for this, the Study Team **recommends over-recruiting case officers** at the start of the trial. Information about past levels of staff turnover in the Member State could be used to estimate the number of case officers expected to leave their posts during the trial. This would minimise the risk that staff turnover results in a lower than required sample size. This would mean that subsequently appointed case officers would not be part of the trial.

2.2.2. Feasibility of randomising applicants

The Study Team concludes that randomising applicants is, in principle, feasible, but that the approach to randomisation would also need to be defined in relation to the specific arrangements in the Member State.

Figure 3 indicates the stages in the process of applying for international protection in Greece. It is the understanding of the Study Team that similar steps are followed in other countries.

The recommendation of the Study Team is that the personnel responsible for scheduling the interview could also undertake two key tasks in a RCT:

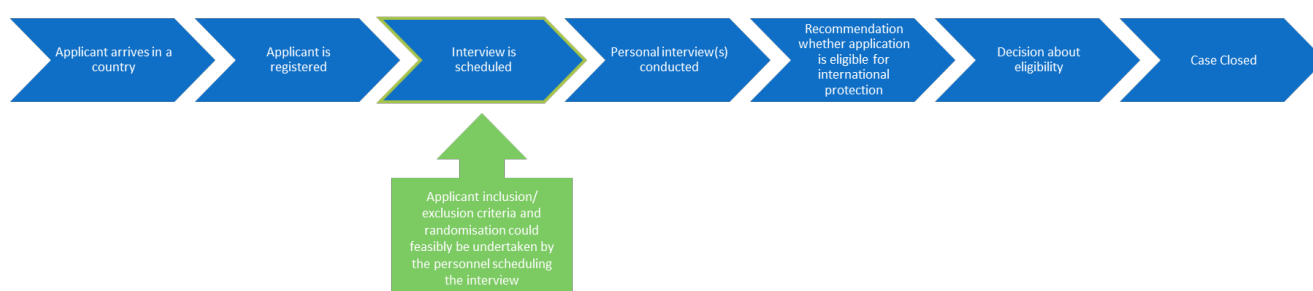
- Excluding any applicants who would not be included in the RCT (for example, applicants under 18 and those identified as vulnerable).
- Randomly assigning applicants to a remote or face-to-face condition.

The ways in which these trial processes are accommodated in a scheduling process will need to be defined in relation to the specific processes in a Member State.

As with randomising case officers, randomisation of applicants could be undertaken using an app, website, or spreadsheet that generates the random allocation. Such tools are commonly used in RCTs.

Randomisation of applicants would be on a rolling basis from a particular point in time.

Figure 3: Where randomisation could fit in the process of applying for international protection



Source: Study Team based on information provided by EUAA staff in Greece

2.2.3. Stratification variables

Stratification refers to dividing the sample into groups (strata) before randomisation. It is undertaken in randomised trials where there is a variable that is known to affect the outcome and ensures that this variable is equally represented in the different arms of the trial. Stratification effectively increases the statistical power of a trial to detect a difference in outcomes between the two groups through forcing balance on the stratification variable.

Case officer experience is a variable likely to affect the primary outcome of a potential RCT (the quality of a personal interview). It is therefore a potential stratification variable, but this would depend on the Member State in question. For example, in Greece all case officers were said to be equally experienced, so stratification might not be needed. If a RCT was implemented in a Member State where not all case officers were equally as experienced, 'experience level' would need to be operationalised into a categorical variable (for example, number of years' experience) and the sample of case officers participating in the trial could be stratified accordingly.

The **physical location of case officers** is another potential stratification variable if an Officer's location means that they are only 'available' to conduct interviews in some locations/some modalities.²⁹ If a RCT was implemented in a Member State in which the location of a case officer imposed limits on their ability to conduct remote or face-to-face interviews, it could be considered as a stratification variable.

Regarding applicants, there are many factors that can affect the quality of an interview (case complexity, for example). In theory, random assignment should mean that any differences in the number of priority and complex cases in each arm of the trial is the result of randomisation and chance, and therefore the associated

²⁹ At the time of writing – February 2023 – in Greece, 12 EUAA case officers are in in Lesvos, 6 in Athens, and 6 in Samos. However, since remote interviewing is used to a great extent in Greece, case officers in all of these locations are able to interview applicants remotely.

uncertainty is captured in the standard error. However, factors such as country of origin or case complexity could be considered as stratification variables.

2.2.4. Randomisation ratio

The randomisation ratio is the proportion of participants who are assigned to one group or another. Although it is typical, and maximises power, to randomise participants in a two-armed trial in a 1:1 ratio (that is, participants have the same probability of being assigned to either arm), this is not a requirement. If there is more capacity to run one form of interview, then the ratio of randomisation could be different to 1:1, so that there is not spare capacity sitting idle. For example, if there are more interview rooms available for face-to-face interviews than for remote interviews, the randomisation ratio could be altered so that slightly more case officers and applicants were assigned to the face-to-face modality.

In Cyprus, for example, remote interviews could be slower than face-to-face interviews as additional administrative work was required to close the file. Ordinarily, Cyprus makes significant use of paper-based systems which often require the case officer to physically sign documents. This could be complicated if interviews were remote. To account for this, the randomisation ratio could assign a larger proportion of cases to face-to-face interviews than to remote interviews, to avoid backlogs.

The randomisation ratio would need to be taken into account when calculating sample sizes (see section 2.4).

It is possible to vary the randomisation ratio during the trial, or to vary the ratio by site, if there are changes in demand or capacity. However, it is not good practice to do so because it means that the probability of assignment to the treatment (or in this case to one of the two interview modalities) becomes correlated with either time, location, or both.³⁰ This is a violation of the central principle of randomisation that all participants in the trial have the same chance of being assigned to each arm of the trial throughout the whole trial. If adjustment of the randomisation ratio was necessary, it should be kept to a minimum, and should be controlled for either through reweighting or controlling for these factors during the analysis.³¹

Box 6: Alternative non-randomised designs

The ToR for this Feasibility and Design Study asked for an investigation of a randomised design.

Randomised trials are a form of 'experimental' design – a research design where the treatment and control groups are (in theory) equal in expectation before the intervention is applied. This is usually achieved through random assignment and allows the evaluator to assume that any change in outcomes is due to the consequence of the intervention as well as chance differences, not any pre-existing characteristics. In analysis, differences that are consequence of chance are quantified through the standard error and probability statements associated with the resultant confidence interval, allowing conclusions to be drawn about the impact of an intervention on a given outcome.

Other designs could be used to compare remote and face-to-face interviews. These other designs would not be as robust as a randomised design as they do not allow causal conclusions to be drawn.

Quasi-experimental designs (QEDs) are used when an experimental design is not feasible because the evaluators are not able to control assignment to experimental groups. QEDs use statistical techniques to create treatment and control groups that are as close as possible to being identical in all respects before the application of the intervention to the treatment group. In the context of the personal interviews in the asylum process, quasi-experimental approaches would involve using data about applicants to 'build' caseloads of similar cases who had received remote and face-to-face interviews and investigate if interview quality differed by modality. Again, since QEDs do not allow unobserved factors affecting interview quality to be controlled for, they are not as robust as a randomised design and would not be able to conclusively show whether face-to-face or remote interviews were equivalent. QEDs are only possible where there are data that allow treatment and control groups to be 'built.' This Study has found that, at the time of writing, the EUAA does not have data that would allow a QED.

Observational designs are those in which the assignment of participants to treatment and comparison groups is not controlled by the evaluator. In the context of the personal interviews in the asylum process, an observational design would not make any changes to the way in which applicants are assigned to remote and face-to-face interviews in a Member State. An evaluator would simply compare interview quality in the two modalities. Statistical techniques could be used to control for factors thought to influence interview quality (for example, case officer

³⁰ Fedorov, V., & Jones, B. (2005). The design of multicentre trials. *Statistical Methods in Medical Research*, 14(3), 205-248.

³¹ For example, this could be done econometrically, as in the analysis of a stepped wedge trial, which involves a similar violation of this assumption. As in a stepped wedge trial, this could be used through the inclusion of time period fixed effects (following Hussey, M. A., & Hughes, J. P. (2007). Design and analysis of stepped wedge cluster randomized trials. *Contemporary clinical trials*, 28(2), 182-191).

experience, case complexity, etc.). However, some of the factors that influence interview quality might be unobservable, and therefore could not be controlled for. This would mean that the results would not be able to conclusively show whether face-to-face or remote interviews were equivalent.

Source: *EEF Glossary*³².

Box 7: The pros and cons of implementing a trial in more than one Member State, and in more than one location within a Member State

A multi-site trial is one in which multiple locations or administrative units are involved in a trial, and the trial is run semi-independently across sites. In a multi-site trial:

- Randomisation is stratified within different sites.
- Different sites can start at different times.
- Data collection may differ slightly between sites.
- Randomisation rates could vary between sites.
- Eligibility criteria could (by exception) vary between sites.
- Analysis is undertaken for all sites.

In the context of a RCT examining the equivalence of remote and face-to-face interviews, there are at least two ways in which a multi-site trial could be conducted:

- A trial in more than one Member State.
- A trial in one Member State, but in a number of different locations within that Member State.

A multi-Member State trial

The Study Team has considered whether it would be feasible and beneficial to implement the potential RCT in more than one Member State at the same time.

The advantages of a trial in more than one Member State are that it would test whether the results were found in more than one location, and as a result be better able to demonstrate replicability and relevance of findings in different settings compared with a one-Member State trial. This might constitute a more convincing evidence base to national stakeholders and to the EUAA.

However, the disadvantages of a multi-Member State trial are the cost and complexity. There is also a risk that (since a multi-site trial aggregate results from all sites) the lack of impact in one Member State might overshadow evidence of impact in another, which could undermine confidence in the result. In addition, a multi-site trial would be an ambitious undertaking without extensive previous experience of commissioning RCTs.

Consequently, whilst the results of a one-Member State trial might not be transferable to another Member State in terms of replicability and comparability, this could be more robust and better implemented with the available resources. Therefore, **the Study Team advises the EUAA to focus on the design and delivery of a robust trial in one Member State at a time, especially in the case of a first trial.** This would allow a narrow focus on ensuring that the remote and face-to-face interview modalities were implemented with fidelity.

A multi-site trial within a Member State

A trial may incorporate many sites within the selected Member State. For example, if there are a number of locations where interviews are held, and case officers are based in those locations, each interview location could be considered as a site.

Source: *Study Team*

2.3. Is it feasible to define clear primary and secondary research questions and to collect data to answer those questions?

A well-designed RCT needs a clear primary research question to be answered by the trial. The trial will be designed (in terms of sample size, power calculations etc.) to answer this primary question.

A trial will commonly have a secondary research question exploring other relevant issues, but which is not necessarily prioritised when making research design decision (i.e., a trial is optimised to answer the primary research question).

For a RCT to be feasible, the research questions need to be matched with available indicators and data.

³² Education Endowment Fund, (n.d). 'Evaluation Glossary: Key evaluation terms and their definitions', <https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/process-and-people/evaluation-glossary>, Accessed: 29/03/2023.

2.3.1. Defining a research question

RCTs can answer different kinds of questions:

- Causal questions: *If I do X, what will happen to Y?*
- Superiority questions: *Is X better than Z at improving Y?*
- Non-inferiority questions: *Is X no worse than Z at improving Y?*

In the case of a RCT of remote and face-to-face interviews, the aim is to answer a 'non-inferiority' question. The Study Team has formulated a **primary impact research question: 'are remote and face-to-face interviews equivalent in quality on average?'** By asking this question, the RCT would test non-inferiority, and thus implied equivalence. The trial would, in other words, provide evidence about whether the two interview modalities were as good as each other – and that remote interviews were no worse.

A RCT will usually also have a secondary question. The **proposed secondary question is: 'are any benefits or costs associated with remote personal interviews compared to face-to-face interviews (e.g., efficiency/cost gains etc.)?'** In an equivalence trial, this second question about costs is very important, since if the modalities are of equivalent quality, then differences in cost might determine the recommendations for practice arising from the trial.

2.3.2. Identifying outcome indicators and data

The starting point for identifying outcome measures is the criteria outlined in Box 8 – these are the ideal characteristics of an outcome measure in a RCT as suggested by the Study Team. In the case of a potential RCT comparing remote and face-to-face interviews, data relevant to the primary research question could come from information commonly collected by national authorities and the EUAA about interview quality.

Box 8: Ideal characteristics of an outcome measure

Interpretable: The meaning of the measure is understood, and it is clear what constitutes a positive or negative change.

Consistent: Collectable in a way that is regular and routine.

Important: Statistically (i.e., the outcome measure allows for statistical assessment of the variable of interest and is predictive of other things that matter) and practically (it relates to concepts that are important).

Common: Cross treatment and control groups, and uncorrelated with levels.

Available: The measure can be collected at reasonable cost and effort.

Source: Study Team

2.3.3. The EUAA Quality Assurance Tool (QAT)

This Study has looked at the possibility of using the QAT developed by the EUAA as the possible outcome measure for the primary research question.

The QAT is an existing measure designed and used by the EUAA and some national authorities. It consists of 14 standards for the personal interview, under which there are 43 indicators. These are divided into the following themes: opening the interview; conducting the interview; substance of the interview; closing the interview; and interview record.³³

For each of the 43 indicators, the assessment can be:

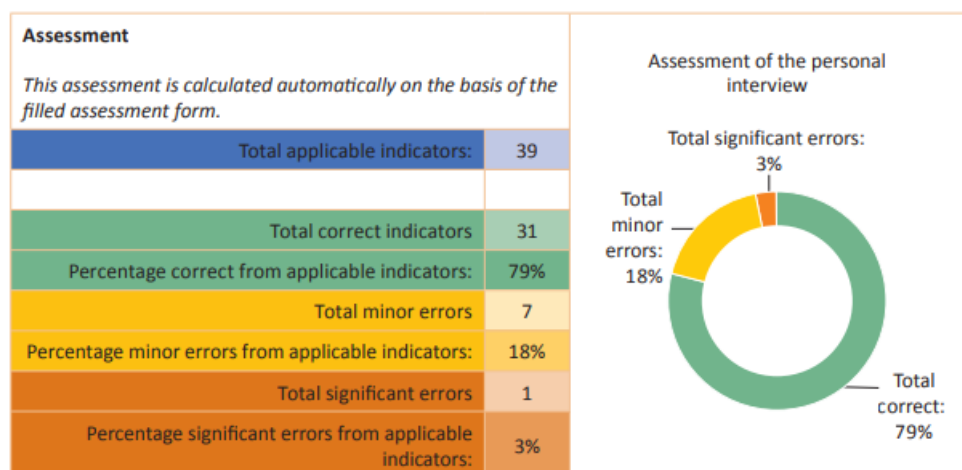
- Correct: the quality requirements are met.
- Minor error: when the error detected does not detract from the overall quality of the interview or decision and would not affect the outcome of the application, AND there are no apparent risks of negative impacts on the applicant, the determining authority or the state.

³³ EASO (2019). 'EASO Guidance on asylum procedure: operational standards and indicators', https://euaa.europa.eu/sites/default/files/publications/Guidance_on_asylum_procedure_operational_standards_and_indicators_EN.pdf, Accessed: 29/03/2023.

- Significant error: when the error detected detracts from the overall quality of the interview or decision and/or may affect the outcome of the application AND there are potential risks or negative impacts on the applicant, the determining authority or the state.
- Not applicable: depending on the national assessment/details if the case.

The output of the QAT in a specific case is illustrated in Figure 4.

Figure 4. Illustration of the output of the QAT 'app' for a single interview



Source: Fictional data for illustrative purposes included in 'EUAA Quality Assurance Tool: Examining the application for international protection' (https://euaa.europa.eu/sites/default/files/publications/2022-01/EASO-Quality-Assurance-Tool-EN_saved.pdf)

Using this output, an overall quality category is assigned to each interview:

- High quality: under 20% minor errors and no significant errors.
- Moderate quality: 20% or more minor errors and no significant errors.
- Low quality: one or more significant errors.

The EUAA have developed an online 'app' version of the QAT which allows a score to be entered for each criterion and generates an overall interview score. In addition, the online version of the QAT provides a structured and qualitative assessment of the interview.³⁴

The QAT is designed to be used at different stages of the personal interview. The QAT can be used:

- While an applicant's case is open: a quality officer, team leader, or line manager of a case officer checks the quality of the interview and instructs the case officer to gather additional information or repeat parts of the interview before making a recommendation.
- For post-hoc assessment of the quality of an interview as part of performance monitoring of the EUAA's support to Member States and to give feedback to case officers as part of their professional development.

Table 3 assesses the QAT against the criteria outlined in Box 8. **The Study Team recommends that the QAT, or a similar tool used in a specific Member State, is a feasible approach to gathering data relevant to the primary research question** outlined in Section 2.3. However, (as explained in Table 3) **further work is needed for the QAT to be ready for use in a trial – the statistical robustness of the measure would need to be investigated.** As part of this further thought is needed about how the output of the QAT could be used in a trial; whether the overall score is an acceptable measure (i.e., a continuous variable), or whether the number of significant and minor errors should be taken into account.

³⁴ At the time of writing – February 2023 – an updated version (version 3) of the app was scheduled for launch with an updated interface intended to be more user friendly.

Table 3. Characteristics of a strong outcome measure of a RCT, how these are exhibited in the EUAA QAT and associated challenges and risks

Ideal characteristic of an outcome measure	How the QAT exhibits the characteristic	Challenges and risks
Interpretable	The QAT has been developed by experts in the field of asylum interviewing; it includes assessment of many different aspects of an interview. A higher quality score is always better. The QAT, as originally developed, provides a structured, qualitative assessment of interview quality, with additional elements, such as the online app, producing supplementary quantitative scores.	The QAT was developed for the purpose of the quality assessment of cases, but it has not been designed or used as an outcome measure for an experimental evaluation. Its statistical robustness has not been examined. The Study Team does not have access to 'historic' scores, ³⁵ so was not able to undertake this statistical testing. Testing statistical robustness would involve understanding its: <ul style="list-style-type: none"> – Reliability: whether it produces the same results in different conditions – Correlation between QAT scores and applicant/case officer characteristics – Inter-rater consistency: how much QAT scores and approaches differ between quality assessors.³⁶ The QAT app produces two outputs: (i) an overall numeric score for the interview (calculated by adding scores for the individual indicators); (ii) whether the interview was of high, medium, or low quality (calculated by the number of significant and minor errors). Before the QAT can be used in a trial, it is necessary to determine which of these outputs would be used in a trial. If using the overall numeric score, a decision is needed about threshold scores that correspond to high, medium, or low quality.
Consistent	The QAT is designed to be used across different countries. There is clear guidance about how to apply it. The QAT is already being used effectively by the EUAA and national authorities in Greece, Cyprus and Malta. ³⁷	
Important	The QAT focuses on interview quality, the key outcome of interest in a possible trial comparing remote and face-to-face interviews.	
Common	The QAT is applicable to both remote and face-to-face interview.	There is less experience in the EUAA of using the QAT to assess remote interviews, but in the view of the Study Team the Tool is suitable for the assessment of both modalities. ³⁸
Available	The EUAA has developed an app which could be used in the trial to facilitate the efficient assessment of interviews.	Quality assessment of personal interviews is a task that requires experience and skill – ideally administered by someone with experience of conducting personal interviews.

³⁵ Historic data means a dataset of closed cases including case characteristics (gender of the applicant, country of origin, age, etc.) and QAT score.

³⁶ Further guidance about how the statistical robustness of the QAT could be tested can be found here: <https://www.testingstandards.net/open-access-files.html> - recommended approach to measuring statistical properties of measurement instruments.

³⁷ The QAT is used differently in different countries in the context of the EUAA operations depending on the needs and the workplans agreed.

³⁸ There were different views among stakeholders about whether the QAT was fully suitable for assessing remote interviews. While the tool was designed primarily with the aim of assessing face-to-face interviews, EUAA staff responsible for the QAT felt it was also suitable for assessment of remote interviews. For instance, the Agency has already issued case officers with practical guidance for remote interviews which covers key areas such as navigating technical issues and improving the transcripts of remote interviews. It also sets out that when problems occur during the interview, this should be recorded in the interview report.

		<p>The length of time taken to assess one interview is not fixed; it depends on the interview length, the case complexity and the experience of the assessor. At a minimum assessing one interview could take 40 minutes for a non-complex claim. At a maximum, it could take several hours, and a complementary interview may be needed. In addition, if there are two quality assessors for each interview, this would require double the time, alongside an additional hour to discuss any differences between assessors and to agree on points. QAT data are not currently available in the format required for a trial: a RCT would require the ability to extract a complete data set including the score for all indicators for all interviews. The reporting functions of the app do not currently allow such a dataset to be generated. Work would be needed by the EUAA to allow data to be extracted from the app in the appropriate format. Discussions with the EUAA suggest that this could be done relatively easily.³⁹</p>
--	--	--

Source: Study Team drawing on interview findings

2.3.4. When could the QAT be used in a RCT?

The Study Team recommends that quality scoring of interviews during a RCT is undertaken based on a written transcript, after the interview is concluded and the case has been closed. While it is acknowledged that some aspects of quality (e.g., body language; room set-up; understanding of the interpreter; the needs of the applicant, such as needing a break; etc.) can only be assessed in person while the interview is being conducted. To allow examination of these factors, a quality assessment could be undertaken 'live' during the interview with the assessor attending the interview as an observer. However, the presence of the assessor in the interview could cause distress to the applicant, could influence the conduct of the case officer, and would incur costs in relation to transport and labour of assessors.

The QAT was designed to be used after an interview has been concluded and a decision has been issued – most of the indicators in the QAT can only be assessed once the interview has been finished. This ex post approach is how quality assessment is currently undertaken in Greece, for example, where it is used most extensively.

This approach requires transcripts to be consistently produced for all interviews included in a RCT. This is not expected to be an issue, since EUAA interviewees explained that most Member States already conduct the quality assessment using a transcript, and the production of a transcript is required by the Directive.⁴⁰

2.3.5. Who would undertake quality assessment in a RCT?

The Study Team suggests that the assessments of interview quality, using the QAT, would be best undertaken by current or ex-case officers who have experience of conducting personal interviews, or by case analysts with experience of the QAT. There is a balance to be achieved between the independence/objectivity of the assessor, and the skill and experience needed to undertake the complex task of applying the QAT to an interview transcript. There was a consensus among consulted stakeholders that it would be very challenging to train a researcher to apply the QAT during a trial. An assessor needs to be familiar with the QAT, understand how to read the interview transcript, and have sufficient knowledge of the relevant national laws and procedures.

Therefore, the Study Team suggests that an acceptable balance of independence and knowledge would be to involve current or ex-case officers or case analysts, *who were not involved in conducting interviews as*

³⁹ In more detail, a proposed solution could involve combining JSON files using external software. Each session-level JSON file is readable into external software (such as Power Bi or the R statistical package). The task is complicated by the need to 'read in' one JSON file per session (i.e., 30 interviews consist of 30 files). However, it should be possible to write a script to read in multiple JSON files and combine them into an internal single dataset. The combined file could be exported in a portable format such as a CSV file.

⁴⁰ Directive 2013/32/EU of the European Parliament and of the Council of 26 June 2013 on common procedures for granting and withdrawing international protection (recast) Article 17.

part of the trial. Stakeholders also mentioned that human resource for quality assessment could be interchangeable between states (i.e., using Greek case officers in a Maltese context etc.), giving the potential for interviews to be quality assessed from another jurisdiction. It is noted that in Greece, interview transcripts are in Greek. Quality assessors would require the relevant language skills.

The Study Team also recommends that all, or a sample of, interviews should be assessed by two assessors as part of the RCT.⁴¹ This is because applying the QAT is, to an extent, a subjective process, and different quality assessors will notice different things during the interview, and so double assessing would provide some mitigation of these. Double assessing a random sample (rather than all) of cases could be a cost-effective option. This could be done as the study progresses to identify any problems in the application of QAT, rather than as a post-hoc quality check.

The interview transcript can be anonymous for the purposes of quality assessment by the assessor. However, the research team undertaking the analysis would need to link the interview quality score to applicants' personal characteristics.

Ideally, quality assessors would not know whether an interview being assessed was remote or face-to-face (i.e., for assessors to be 'blind')⁴². This is to reduce the risk that the assessor is consciously or unconsciously biased and so undertakes the quality assessment in a different way for remote and face-to-face interviews.

It is likely to be difficult to blind assessors to whether the interview being assessed was remote or face-to-face, since indicators of the modality would likely be apparent in the transcript. It might be possible for transcripts to be redacted in advance of quality assessment, to remove references to the modality, but this would involve time and cost.

2.3.6. What other data are needed for a trial?

For the statistical analysis, the researchers conducting the trial would need access to:

- Full data set from the QAT (as explained above). The fields from the QAT needed for a trial are set out in Annex 2.
- Full data set about applicant characteristics and case characteristics. Annex 3 lists the fields needed, using the dataset used in Greece as an example.
- Characteristics of the case officers: years of experience, qualifications.

The researchers would need to link these different data. This would require a common identifier in all data sets to facilitate data linkage, so that the applicant, their characteristics, the QAT score of their interview and the identity and characteristics of the case officer can all be put together in a single data set. EUAA interviewees consulted during this study have indicated that while full applicant details are not currently linked to QAT score data at present, it would be possible to link the two data sets based on a common or case identifier, with a need to remove certain identifying fields such as names.

⁴¹ The assignment of assessors would need to be worked out by the contractor and reviewed in the pilot.

⁴² Blinding is where information about the assignment of participants to their experimental group (e.g., control or treatment) is concealed from the evaluator, the participants, or other people involved in the study until it is complete. Blinding can be introduced at various points in an evaluation, for example: randomisation (the person carrying out the randomisation does not know any information that could be used to identify the participants being randomised); analysis (the person carrying out the analysis does not know any information that could be used to find out which participants are in which experimental group); measuring outcomes (the person administering/measuring outcomes does not know whether participants are in the treatment or control group). Failure to blind can introduce bias – not necessarily intentionally, but those involved might subconsciously behave differently if they know that a participant is receiving an intervention/control. See: <https://educationendowmentfoundation.org.uk/projects-and-evaluation/evaluation/process-and-people/evaluation-glossary>

2.4. Is the number of interviews conducted likely to generate a sufficient sample size within a reasonable timeframe?

It is important to consider what sample size the RCT will need to produce statistically significant effect sizes to confidently test equivalence between the intervention and 'business as usual'. Based on this, an estimate of the length of the trial can be arrived at.

Power calculations are conducted prior to running a randomised trial. Power calculations can:

- Indicate how large a trial needs to be to detect a particular effect size, known as the minimum effect size of interest (MESI) or
- Indicate the size of effect the trial can detect with reasonable probability, should such an effect exist, known as the minimum detectable effect size (MDES).

Calculating the MESI is primarily useful when there is a minimum effect that is needed for an intervention to be worthwhile. For example, if an intervention costs €5,000 per person to administer, and aims to prevent an outcome that costs €100,000 per person affected, then any effect smaller than 0.05 (a five percentage point reduction in the probability that the outcome occurs, per person treated), would not be of interest, and the trial does not need to be designed to detect effects smaller than this.

A trial comparing remote and face-to-face interviews would be an inferiority trial. In order to undertake power calculations for such a trial, an **inferiority threshold (or margin of equivalence)** would need to be established. This is akin to a MESI. **The inferiority threshold in this trial would be the acceptable difference in quality between remote and face-to-face interviews.** The inferiority threshold might be zero – i.e., no detectable difference in quality is acceptable.

Defining the inferiority threshold would need to be done in close consultation with the national authorities in the Member State in which a trial is to be implemented and the EUAA. Legal responsibilities in national law and the Directive would need to be considered.

To define the inferiority threshold, the output from the QAT would also need to be determined.

As outlined in Table 3, further work is needed to operationalise the outputs of the QAT app so that they can be used in a trial. If the primary outcome measure of the trial was the overall numeric quality score for an interview from the QAT (i.e., a continuous variable), then the inferiority threshold could be expressed in terms of the acceptable percentage difference, if any, between the score of a remote interview compared to a face-to-face interview. If the primary outcome measure of the trial was a binary measure, then the inferiority threshold could be expressed in terms of the acceptable difference, if any, in the percentage of interviews assessed as 'low quality' in each arm or the trial. Unfortunately, the information needed to calculate sample sizes is not available to the Study Team. The information needed to calculate the sample size in an inferiority trial is:

- Whether the outcome measure is a continuous variable (i.e., a numeric quality score), a categorical variable (i.e., high, low, or medium quality) or a binary variable (i.e., acceptable quality/unacceptable quality).
- The statistical properties of the QAT such as mean scores and standard deviation – and, if possible, how much interview quality varies by applicant and case officer characteristics.
- The inferiority threshold for the trial, agreed with the Member State national authority.
- The number of case officers who could participate in a trial, in a Member State.
- The number of interviews conducted per week per case officer, in the Member State.
- The inter-cluster correlation rate (comparing the variance within clusters with the variance between clusters).

The output of the power calculations would be: the number of case officers needed to participate in the trial and the number of interviews each case officer would need to conduct in total over the trial in order to be able to robustly assess non-inferiority at the agreed threshold.

It is recommended that the EUAA undertakes an analysis of the statistical properties of the QAT. Through doing this, the Agency would be able to obtain many of the pieces of information listed above and undertake sample size calculations.

2.5. Is it feasible to implement a RCT ethically?

A trial comparing remote and face-to-face interviews raises important ethical considerations, with potential harms to stakeholders. If a trial cannot be designed which complies with ethical standards, then the trial would not be feasible.

For some stakeholders, the concept of deploying randomisation in the asylum system raises ethical questions. In the context of this possible RCT, there might be concerns that applicants assigned to the remote condition will be disadvantaged in some way or have a poorer experience.

The ethical issues raised by evaluations employing randomised designs can normally be managed and mitigated through the way in which the trial is conducted and monitored.

2.5.1. Risk of harm to participants

When considering the risk of harm to applicants and case officers, the following are relevant to the overall assessment of the ethics of a trial comparing remote and face-to-face interviews:

- **The extent to which remote interviews are 'business as usual' in a given Member State.** As indicated in Box 1, in a Member State already using remote interviews (such as Greece), a trial would be evaluating existing practice which, it is assumed, has built-in safeguards and procedures for applicants and case officers.
- **The evidence that remote interviews and face-to-face interviews can both be of good quality.** In the context of designing randomised trials, the concept of 'equipoise' refers to the situation in which there is uncertainty about whether one 'treatment' is better than the other. This is the case in relation to remote and face-to-face interviews – there is evidence that both can provide a good quality interview. The Study Team is of the view that it is ethically acceptable to randomise applicants to either remote or face-to-face modalities, since there is a plausible case that these modalities are equivalent in terms of the ability to conduct a good quality interview. There is no evidence that either modality is more disadvantageous.
- **The extent to which remote and face-to-face interviews can be implemented with fidelity in a given setting.** If implemented correctly, this implies following operating procedures and guidelines to ensure applicants are protected and minimises the risk of harm.
- **Exclusion of vulnerable applicants.** As discussed in Section 2.1.9 it is recommended that vulnerable applicants are excluded from the trial.

2.5.2. Informed consent

The extent to which applicants and case officers are asked to give informed consent is an important decision to be taken in the design of the trial comparing remote and face-to-face interviews. In some trials, participants are asked to give informed consent for their involvement. For example, patients in a trial which tests a drug will be asked to consent to the trial, which involves consenting to be randomly assigned to the treatment or control group (they will not know which). Trials comparing two 'business as usual' modalities can be conducted in line with ethical standards even when informed consent is not sought. For example, currently, in Greece, applicants can be given a remote or face-to-face interview, with the decision about the modality made by the person scheduling the interview. What would be different if a trial was implemented in Greece is that the assignment to a remote or face-to-face modality would not be made by the person scheduling the interview, it would be decided randomly – using a randomiser tool.

The practicalities of seeking informed consent in a trial comparing remote and face-to-face interviews also need to be considered. This is the question of who would administer consent, at what part in the process outlined in Figure 4. In Greece, for example, applicants are notified of the date, time, location, and modality of their interview via letter or email, not via a face-to-face conversation with the scheduler. If informed consent was to be required, then an additional step would need to be inserted into the process to administer consent.

It is the view of the Study Team that it would be ethically acceptable not to seek the informed consent of applicants to be randomised to an interview modality where a Member State already uses remote interviewing. In a Member State that was implementing remote interviews for the first time for the purposes of the trial, there might be a stronger case for requiring informed consent, because the trial is having a more significant impact on applicants. Much would depend on the legal basis in the Member State and the extent to which both interview modalities were implemented in line with EUAA guidance.

The ethical review board from which approval for the trial is sought may have their own rules and approaches to seeking informed consent.

2.5.3. Monitoring adverse events and harm

During a trial, as is good practice in RCTs, real-time monitoring should be undertaken in order to detect any unintended consequences of the trial. Monitoring might include, for example, determining whether being in the trial, or in one arm of the trial, is associated with any changes in the following:

- Length of time between registration and decision
- Length and number of interviews per applicant
- Rate of rejection/approval
- Appeals
- Complaints
- Applications withdrawn before the interview takes place
- Non-attendance at interview

Procedures should also be in place to allow those involved in the delivery of remote and face-to-face interviews during the trial to report any concerns.

As part of planning a trial that involves risks, it is common to specify types of harms or incidents that would result in pausing the trial, and to specify how and by whom they would be investigated.

2.5.4. Confidentiality and information sharing

It would be necessary in a trial for data about applicant characteristics, case officer characteristics, and interview quality scores to be shared with the contractor undertaking the analysis.

The contractor would need to put in place a **data protection plan** which protects the confidentiality of information.

While the Study Team does not believe that seeking informed consent from applicants for inclusion in the trial is necessary, applicants could be asked at the start of the interview for consent for their data to be shared with a contractor for research purposes. This could be included in the existing preamble used at the start of an interview.

When making arrangements for confidentiality and information sharing, the legal framework of the Member State should be assessed and also the view of EUAA legal services, to ensure compliance with national law and the Directive.

2.5.5. Ethical approval

A trial would require approval from a recognised ethical review board. Applying for ethical approval would be the responsibility of the contractor undertaking the trial.

2.6. Is it feasible to implement a RCT in social and political contexts?

The political, social, and cultural sensitivities relating to asylum are also important considerations in relation to the feasibility of a trial. This is in addition to practical considerations about whether remote and face-to-face interviews can be delivered with fidelity (Section 2.1), and whether randomisation can be accommodated in national processes (Section 2.2). While social and political factors are not usually considered as part of feasibility studies, it is clear from the data collected to inform this Study that issues related to asylum are highly sensitive. Consequently, this makes stakeholders wary about the ethics of a trial (see Section 2.5) and how it will be perceived. These factors might affect, for example, the extent to which national authorities can support a trial and how stakeholders such as lawyers, advocacy organisations, and the public perceive a trial. These factors could affect the implementation of the trial (for instance, if case officers feel uncomfortable delivering interviews as part of a trial this might diminish their effectiveness as interviewers).

Importantly, **full support of national authorities** for a trial is essential and, linked to this, **full understanding in the national authority that the trial would involve in random assignment** of applicants and case officers to remote and face-to-face modalities, which would to some extent limit the discretion of those scheduling interviews.

There may also be a need to consider reputational risk for the EUAA and Member States if complaints are raised by participants in the RCT, or other stakeholders, as well as any management or mitigation strategies that could be utilised.

A pilot phase of a trial (which the Study Team recommends) could also be used to learn about these social and political aspects of feasibility.

2.7. The value of a pilot phase of a trial

Study Team recommends that a pilot phase is undertaken at the start of a trial. A pilot phase consists of a period of time in which to test trial processes, collect and undertake analysis of data about applicants, case officers, and interview quality. It would provide the contractor implementing the trial with an opportunity to make changes before the trial is fully set up. The aims of the pilot would include:

- Identifying 'real-world' logistical challenges regarding implementation of the interviews (for example, scheduling delays, issues of transporting applicant or case workers, etc.)
- Identifying any issues regarding trial processes relating to applying eligibility criteria, randomisation, etc.
- Collecting and analysing data about a small number of interviews to inform decisions about whether additional stratification is needed or whether the randomisation ratio needs to be altered.
- Allowing ethical risks and issues arising from the assignment of interviews to remote or face-to-face modalities to be identified and be acted on quickly.
- Allowing social and political aspects of feasibility to be tested.

It is recommended that the statistical properties of the QAT are investigated before a pilot.

A short report could be produced as an output of the pilot phase, outlining findings from the pilot and listing any recommendations on ways in which the trial should be amended before full implementation.

2.8. The need for implementation and process, and cost evaluation

An implementation and process evaluation provides insight into the 'how and why' of the results from the impact evaluation which is useful for understanding how the interview process might be improved.

The cost evaluation would provide essential information about the cost – in terms of staff time, materials, and other resources – of an average remote interview and an average face-to-face interview.

2.8.1. Questions answered by an implementation and process evaluation

An implementation and process evaluation provides **insight into whether the key 'ingredients' of the intervention were delivered as planned** (with fidelity). For example, whether rooms were suitable, internet connections were sufficient, case officers were trained, etc. Questions addressed by an implementation and process evaluation could include:

- To what extent were remote and face-to-face interviews implemented with fidelity?
- To what extent were any adaptations made to the implementation model for remote and face-to-face interviews? Why?
- What were the barriers and facilitators to implementing remote and face-to-face interviews with fidelity?
- Was there any evidence of contamination?

The **experience of applicants, case officers, interpreters, and lawyers** is not captured in the QAT (the proposed primary outcome measure) but could usefully be explored in an implementation and process evaluation as a route to detecting unintended consequences, exploring fidelity, and understanding the full consequences of interview modality.

- How were remote and face-to-face interviews experienced by interview participants (applicants, case officers, interpreters, lawyers)?
- What strengths and weaknesses of remote and face-to-face interviews are perceived by interview participants (applicants, case officers, interpreters, lawyers)?

An implementation and process evaluation could also **capture the views of the relevant national authority** about hosting the trial and about the use of remote interviews.

- How were remote and face-to-face interviews experienced by the national authority?

- What strengths and weaknesses of remote and face-to-face interviews are perceived by national authorities?

An implementation and process evaluation should probe to understand if there are any **positive or negative unintended consequences** to remote interviews compared to face-to-face interviews. These might be effects that would not be detected through, for example, analysis of management data of the quality of interviews as measured by the QAT but might be only detected qualitatively.

- To what extent, if any, does the use of remote interviews have positive or negative unintended consequences for applicants, interpreters, case officers, national authorities, and the EUAA?

2.8.2. Data collection approaches for the implementation and process evaluation

The methods for an implementation and process evaluation should be proposed by the contractor implementing the RCT. It is important that data are collected from all relevant stakeholders. The methods could include observations of a small number of interviews (in both modalities) by researchers, to develop a richer assessment of the interview, beyond what can be ascertained from a quality review based on a transcript. Interviews are audio recorded in Greece and Malta, so listening back could be a possibility, in addition, or as an alternative, to shadowing.

Analysis methods should be proposed by the contractor implementing the evaluation.

2.8.3. Questions answered by a cost evaluation

As part of a trial, it will be important to understand the cost of a remote compared to a face-to-face interview. This is important because if the quality is found to be equivalent across the modalities, cost could be a deciding factor on use of remote or face-to-face interviews in the future.

- What are the monetisable costs involved in scheduling, conducting (including interview duration, the need for re-scheduling, case worker commuting time) and concluding a remote and face-to-face interview?
- What is the monetised, average cost of implementing a remote interview compared to a face-to-face interview?
- Who bears these costs?

Cost data held by national authorities (or the EUAA) would be required to understand the costs. It might be useful to use categories for costs:

- Staff costs – national authorities
- Staff costs – interpreters
- Buildings and facilities costs (capital expenditure)
- Materials and equipment
- Miscellaneous

Analysis methods should be proposed by the contractor implementing the evaluation. The analysis would ideally result in a unit cost per interview for each modality.

3. DESIGN OF A POTENTIAL RCT

One of the tasks specified for this Study was to 'formulate a robust and statistically and ethically sound RCT design framework including an implementation plan'. However, because **much of the detailed design of a RCT would be specific to the Member State in which the trial was implemented**, and because historic data about interview quality linked to applicant characteristics are not available, the Study is not able to fully achieve this.

Therefore, this section outlines the kind of information which should be included in a future trial protocol⁴³ to be prepared by a contractor planning a trial in a specific Member State. This chapter includes recommendations, where possible, for the RCT design where the Study has allowed such design recommendations to be reached. This chapter could be used as a resource by a contractor appointed to plan and conduct a RCT in a Member State.

Some of the information that should be contained in a protocol has already been outlined in earlier parts of this report. To avoid repetition, readers are referred to the relevant sections.

3.1. Study rationale and background

This section of the protocol should describe what is being evaluated, and why.

The rationale for a study comparing remote and face-to-face interviews is to build the evidence base in order to effectively achieve the goals of the CEAS.

This section of the protocol could include other specific motivations for the trial, relevant to the Member State in which the trial was to be implemented. It should contain the background, EU and national policy context, similar to the information in Section 1.1 of this report.

3.2. Intervention

This section of the protocol should describe the intervention in sufficient detail to allow replication, including the stakeholders involved. The programme theory should be explained, along with process maps etc.

This section should describe the personal interview and its role in the asylum process (as outlined in Section 1.1 of this report), specific to the Member State in which a trial was to be implemented. The recommendation of this Study is that the intervention is defined as follows:

- Personal interviews to determine eligibility for international protection.
- Remote interviews are those in which the interviewing case officer and applicant are not in the same room.
- Face-to-face interviews are those in which the interviewing case officer and applicant are in the same room.
- The location of interpreters and lawyers is not central to these definitions.

This section of the protocol should include a description of all the stakeholders involved in the delivery of interviews. The descriptions in Section 2.1.3 of this report could be used as a starting point for national-specific definitions.

The programme theory of personal interviews should be further elaborated by the contractor implementing the trial. A draft is outlined in Box 2 in Section 2.1.1, and this could be tailored to the national context.

3.3. Impact evaluation design

This section of the protocol should describe:

- Research questions.
- Outcome measures.

⁴³ It is an internationally recognised good practice for a RCT to have a clearly defined protocol, prepared in advance of the trial. When completed, a protocol records all aspects of the design, implementation, and analysis plans for a RCT.

- Design/type of trial (e.g., parallel group, crossover, factorial, single group), allocation ratio, stratification, and framework (e.g., superiority, equivalence, non-inferiority, exploratory); time schedule of enrolment (i.e., how many cases would be allocated to treatment and control per week over the life of the trial).
- Participants, including eligibility criteria and exclusion criteria.
- Sample size and power calculations.
- Data collection plans: plans for assessment and collection of outcome, baseline, and other trial data, (including processes to promote data quality – e.g., duplicate measurements, training of assessors – and a description of study instruments along with their reliability and validity, if known; plans for data entry, coding, security, and storage.
- Analysis plans – the methods for analysing the data collected.

3.3.1. Aim and research question

This Study has recommended the following:

- Primary impact research question: 'are remote and face-to-face interviews equivalent in quality on average?' (non-inferiority and thus implied equivalence)
- Secondary impact question: 'are any benefits or costs associated with remote personal interviews compared to face-to-face interviews (e.g., efficiency/cost gains etc.)?'.

3.3.2. Outcome measures

This Study has recommended the primary outcome of interview quality as measured by the EUAA QAT (or similar tool used within the Member State).

This section of the protocol should include a full description of the QAT (or similar national tool) including its strengths and limitations. See Section 2.3 of this report for a description of the QAT.

This Study has recommended that, since the QAT has not been used for an experimental evaluation, further work would be needed (ideally before a pilot phase of a trial) to investigate the statistical properties of the QAT and decide whether the categorical (high, medium, low quality) or continuous (numeric quality score) variable should be used as the outcome measure. The use of the QAT should then also be tested during a pilot phase of the trial. This part of the protocol is likely to be finalised after the analysis of data from the pilot.

3.3.3. Design/type of trial and approach to randomisation

The design of the trial which this Study has recommended is a cluster RCT with randomisation at the level of the case officer to form the clusters, as well as randomisation at the level of the applicant to form cases. It is a trial testing equivalence or non-inferiority.

This design is proposed because it should allow the many factors of a case which could affect interview quality to be controlled for, thus allowing for the best possible approach to generate robust evidence of the equivalence/non-inferiority of the modalities.

The application of the inclusion criteria and randomisation would be undertaken by the national authority responsible for scheduling the interview. It is recommended that a randomiser tool is used.

This Study has not identified any stratification variables. However, piloting the randomisation might result in a decision to stratify (see Section 2.2.3).

The protocol should state the randomisation ratio, which would be determined in the particular Member State context, following a pilot of the randomisation (see Section 2.2.4).

3.3.4. Participants/sample

This Study has recommended that the main study participants are applicants applying for international protection eligibility in a given Member State, and that the following would not be included in a trial:

- Applicants identified as vulnerable (according to the criteria applied in the Member State in question).
- Applicants under 18.

The protocol should provide other details specific to the Member State in which a trial is implemented, for example, the locations of the applicants.

Any other inclusion or exclusion criteria specific to the Member State should be included here.

3.3.5. Sample size and effect size/power calculations

See discussion in Section 2.4 of this report. This Study can provide information on what information is needed to calculate power and sample size. This part of the protocol can be finalised after analysis of historic QAT data has been undertaken and an inferiority threshold for the trial has been agreed.

3.3.6. Data collection

This part of the protocol should specify how data from the QAT (or similar national tool), from the operational data collection (or similar national case management information), and about case officers will be obtained, how it will be transferred, stored, cleaned, and coded. Annexes 2 and 3 list the information fields that would likely be needed to be shared with a contractor delivering the trial.

Information about the identity of quality assessors could be included here, along with how they will be trained to undertake quality assessment, how their work will be monitored, arrangements for double coding (e.g., having two assessors undertake quality assessment of a random subset of interviews) to check inter-rater reliability. See Section 2.3.5 of this report for more details.

3.3.7. Analysis

A detailed analysis plan would need to be prepared by the contractor implementing the trial. This part of the protocol would be finalised based on the nature of the trial (i.e., as a non-inferiority trial), analysis of the statistical properties of the QAT, and data from the pilot.

Table 4 below briefly outlines the type of analytical approach that would be taken according to the variable type of the (QAT) outcome measure.

Table 4 – Possible approaches to analysis by type of variable

	Binary variable	Continuous variable	Categorical variable/multiple categories
Possible approach to analysis	Logistic or probit regression using hierarchical model or clustering on case officer. The approach not chosen, or a linear probability model, could be used for a robustness check.	Ordinary least squares (OLS) linear regression with clustering on case officer (hierarchical mixed model could be alternative).	Ordinal logistic regression using hierarchical model with clustering on case officer is an alternative, with likely more easily interpretable results).

3.4. Implementation, process and cost evaluation

This section of the protocol should describe:

- Research questions
- Participants and sampling
- Data collection plans
- Analysis plans

3.4.1. Aim and research question

The implementation, process and cost evaluation would aim to:

- Monitor fidelity.
- Explore the experiences of applicants, case officers, national officials, interpreters.
- Detect positive/negative unintended consequences.
- Estimate cost of the modalities and factors that affect cost.

3.4.1. Participants and sampling

A list of the expected numbers of participants consulted using different data collection methods, and how they would be selected.

3.4.2. Data collection

This could involve interviews and online surveys with applicants, case officers, etc. as well as observing interviews in both modalities and/or reviewing audio recordings of interviews in both modalities – to supplement the quality assessment based on the interview transcript.

The sources of cost data should be specified; for example, staff costs (wages) could be requested from national authorities, secondary sources could be used for inflation rates, etc.

3.4.3. Analysis

This part of the protocol should explain how the (mainly) qualitative data collected in the implementation and process evaluation would be analysed; for example, thematic coding.

For the cost evaluation, this part of the protocol should specify how the costs of the interviews will be estimated, what assumptions will be used, and what the output would be.

3.5. Ethics and data protection

This section of the protocol sets out all the possible ethical risks and how they will be mitigated, including considerations around safeguarding of participants and researchers. It should include how ethical approval will be sought.

Data protection arrangements should be described in detail, including a data flow diagram, clear indications of data owners, processors, etc.

The contractor implementing the evaluation would need to propose a data protection plan for all elements of the evaluation, specific to the Member State and the data being collected, in compliance with the General Data Protection Regulation (GDPR) and national law. See Section 2.5 of this report.

3.6. Risks

This part of the protocol should specify all risks to the trial – including data collection and analytical risks.

This Study has identified the following potential risks which could be included here:

- **Implementation/fidelity risk** – including whether the Member State in question can implement remote and face-to-face interviews in line with guidance, including having sufficient numbers of trained case officers, appropriate interview rooms, etc.
- **Risk to trial processes** – there is a risk that the assessment of eligibility and randomisation, which would be conducted by those responsible for scheduling interviews, is not undertaken consistently. Schedulers are likely already under time and resource related pressures. Training will be required to avoid this scenario.
- **Outcome measure risk** – the QAT is considered by stakeholders to be a good measure of interview quality, but it was not designed for use in a trial, so its statistical validity is not known; the use of the QAT relies on the availability of trained quality assessors. Additionally, there is a risk that quality assessors (consciously or unconsciously) assess one interview modality more harshly than the other, biasing the results.

- **Data access risk** – the data needed for the trial are owned (depending on the national situation) by the EUAA and national authorities. Data sharing agreements need to be put in place, likely requiring the involvement of data protection experts. This could be costly and create delays.
- **Ethical risks** – these relate to the inherent vulnerability of applicants, and the need to ensure that any risk of harm to applicants is minimised. Decisions about whether to seek informed consent and/or consent to data sharing need to be taken.
- **Analytical risks** – there might be uncertainties related to the analysis plan (for example, missing data).

3.7. Timeline

This part of the protocol should provide a detailed timeline for all stages of the trial.

The contractor should specify:

- Timeline for implementation and analysis of a pilot.
- Period in which the trial would be implemented.
- Points when data will be collected – including in the implementation and process, and cost evaluations.
- Time period needed for analysis.
- Dates of preliminary results, peer review, and final results.

4. CONCLUSION AND NEXT STEPS

4.1. Key conclusions about feasibility

This study assessed the feasibility of implementing a RCT comparing face-to-face and remote interviews. The aim of such a RCT would be to conduct an evaluation of the impact of these different interview modalities on the quality of the interview using the following criteria:

- Is it feasible to define the 'intervention' and implement it with fidelity in operational contexts?
- Are there feasible options for randomisation?
- Is it feasible to define a clear primary and secondary research question and to collect data to answer those questions?
- Is the number of interviews conducted likely to generate a sufficient sample size within a reasonable timeframe?
- Is it feasible to implement a RCT ethically?
- Is it feasible to implement a RCT in social and political contexts?

Overall, the Study finds it is, in principle, feasible to implement a RCT, although a final assessment of feasibility would need to be specific to a Member State in which a trial was to be implemented.

4.2. Key conclusions about design

This Study has set out a suggested design for a trial with the following characteristics:

- A cluster RCT with randomisation at the level of the case officer to form the clusters, as well as randomisation at the level of the applicant to form cases.
- Testing equivalence or non-inferiority.
- Using interview quality as the key outcome measure, as assessed by the existing EUAA QAT.
- Implemented in one Member State (rather than a multi-Member State/site trial).

4.3. Next steps for the EUAA and contractor

The assessment of feasibility and proposed design set out in this Study is limited because many details of a trial (impact evaluation) depend on the national context in which a trial was to be implemented. If the EUAA or a national authority wanted to pursue a trial, Table 5 outlines possible next steps, assuming the EUAA is commissioning the trial.

Action	Owner
Undertake an analysis of the statistical properties of the QAT (either within the Agency, or by appointing an external contractor).	EUAA
Select a Member State that wants to implement a trial (impact evaluation), and secure national approval and support. Consider signing a Memorandum of Understanding setting out how each party will support and facilitate the trial.	EUAA
Appoint a contractor to undertake the pilot and main-stage trial (with a break clause in the contract if the pilot highlights issues which make the trial no longer feasible).	EUAA
Map, in detail, the processes and procedures involved in scheduling and conducting interviews in that Member State, collect data about the number of cases and case officers who would be involved etc. Based on this, undertake initial power calculations, and highlight where investment would be needed (if any) to implement remote and face-to-face interviews with fidelity.	Contractor in partnership with EUAA and National Authorities
Co-design and prepare a draft study protocol.	Contractor in partnership with EUAA and national authority

Put in place data sharing and protection arrangements, fully agreed with the national authority.	Contractor
Peer review of the protocol	EUAA
Ethics approval	Contractor
Communications planning	EUAA in partnership with national authority and contractor
Prepare statistical analysis plan	Contractor
Pilot	Contractor in partnership with EUAA and national authority
Analysis of pilot data, adjust protocol and finalise statistical analysis plan	Contractor
Peer review of protocol and statistical analysis plan	EUAA
Implementation of trial (impact evaluation)	Contractor

Source: Study Team

ANNEXES

ANNEX 1: Data collection activities for the Feasibility and Design Study

The table below outlines the data collection activities undertaken to inform this Feasibility and Design Study

Data Collection Activity	Dates/Study Phase	Details
Scoping Interviews	10.08.2022, Phase 1	Data Analysis and Research Sector, C3
	11.08.2022, Phase 1	Operational Programming Sector, C1
	12.08.2022, Phase 1	Asylum Processes Sector, C3
	19.08.2022, Phase 1	Training and Learning Research and Analysis Sector, C2
	30.08.2022, Phase 1	Operational Support Centre, C1
Project management interviews	14.09.2022, Phase 2	Call with Project management from Cyprus, Greece, and Malta operations
	27.09.2022, Phase 2	Cyprus Project management meeting
	28.09.2022, Phase 2	Malta Project management meeting
	06.10.2022, Phase 2	Greece Project management meeting
Additional meetings/Interviews	26.10.2022, Phase 2	Operational Support Centre, C1
	26.10.2022, Phase 2	Asylum Processes Sector, C3
	03.11.2022, Phase 2	Operational Quality Procedures and Tools Sector, C1
	17.11.2022, Phase 3	Operational Support Centre, C1 EUAA Management, Greece and Malta operations
	29.11.2022, Phase 3	Quality Management and Evaluation Sector, C4 Greece operations
	13.12.2022, Phase 3	EUAA staff demonstrated the EUAA QAT including how the online app worked
	5.01.2023, Phase 3	Interview with Greece operations to understand the details of how personal interviews are scheduled in Greece
Key Documents/Desk Research/Literature ⁴⁴	Phase 1,2	EUAA QAT: https://euaa.europa.eu/sites/default/files/publications/2022-01/EASO-Quality-Assurance-Tool-EN_saved.pdf
	Phase 1,2	Practical recommendations on conducting the personal interview remotely: https://euaa.europa.eu/sites/default/files/easo-practical-recommendations-conducting-personal-interview-remotely-EN.pdf
	Phase 1,2	EASO Practical Guide: Personal Interview: https://euaa.europa.eu/sites/default/files/public/EASO-Practical-Guide-Personal-Interview-EN.pdf
	Phase 1,2	EUAA operational plan to Greece 2022-2024: https://euaa.europa.eu/sites/default/files/EL_OP_Greece_2022-2024_signed.pdf

⁴⁴ This includes only the key documentation, desk research, and data sources utilised at this stage of the Study rather than a comprehensive list of all the sources that the Study Team has reviewed from the Agency throughout the Study.

	Phase 1,2	EUAA operational plan to Malta 2022-2024: https://euaa.europa.eu/sites/default/files/EUAA_2022-2024_OperationalPlanToMaltaAmendment_0.pdf
	Phase 1,2	EUAA operational plan to Cyprus 2022-2024: https://euaa.europa.eu/sites/default/files/OP_CY_2022-2024.FINAL_.pdf
	Phase 2	Guidance for using pilot studies to inform the design of intervention trials with continuous outcomes: https://pubmed.ncbi.nlm.nih.gov/29403314/
	Phase 2	Directive 2013/33/EU of the European Parliament and of the Council of 26 June 2013 laying down standards for the reception of applicants for international protection (recast): https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32013L0033
Data/Statistics	Phase 2	Internal data from EUAA relating to case numbers, numbers of case officers, numbers of interviews in Cyprus, Greece, and Malta in 2022
Capacity-building workshops	13.10.2022, Phase 2 16.02.2023, Phase 3	Ensured that Agency staff have insight into how the Feasibility and Design Study has been undertaken by the Study Team; understand the different aspects of the Feasibility and Design Study, why and how decisions were made about these different aspects; understand the pros and cons, risks, mitigations, and trade-offs which will be analysed as part of the Feasibility and Design Study. These have been included as a data collection activity within the Study because discussion at the workshops provided additional insight and highlighted issues for consideration.
Written responses	January 2023, Phase 3	Written response from EUAA staff based in Greece about scheduling processes and other practices.

Source: Study Team

ANNEX 2: Data fields needed from the QAT, for all interviews included in the trial

Field	Sub-field
Case file reference	Interviewer
	Team/unit
Applicant	Country of origin
	Sex
	Age
	Special Needs
Case data	Date of lodging the application
	Date of interview
	Interview conducted through interpreter
	Language of the interview
	Legal representative present during the interview
	Legal representative present during the interview
	Grounds for the application
	Decision outcome
Assessment	quality assessor
	assessment date
	assessment based on
Other	additional info
Assessment form Drop down list for each criterion Correct; minor error; significant error; not applicable	1.1
	2.1
	2.2
	2.3
	2.4
	2.5
	2.6
	3.1
	4.1
	4.2
	5.1
	5.2
	5.3
	5.4
	6.1
	6.2
	6.3
	6.4
	6.5
	7.1
	7.2

	7.3
	7.4
	7.5
	8.1
	8.2
	8.3
	8.4
	8.5
	9.1
	9.2
	10.1
	11.1
	11.2
	12.1
	12.2
	12.3
	13.1
	13.2
	13.3
	14.1
	14.2
	14.3
Conclusion	The interview allows an effective and correct decision to be made
Assessment	Total applicable
	Total correct
	% correct from applicable
	Total minor errors
	% minor errors from applicable
	Total significant errors
	% significant errors from applicable
	Overall quality

Source: EUAA QAT application⁴⁵ and EUAA guidance on asylum procedure⁴⁶

⁴⁵ <https://euaa.europa.eu/easo-updated-quality-assurance-tool-application-now-available>

⁴⁶ <https://euaa.europa.eu/publications/guidance-asylum-procedure>

ANNEX 3: Indicative data fields needed from relevant data sources relating to applicant/case characteristics.

The table below is intended to provide a guide for future contractors implementing a RCT about the kinds of information held by the EUAA. It provides an indication of the kinds of data which could be shared as part of a trial to provide essential information about applicant demographics, details of their case, outcomes etc.

Table 6: Fields relating to applicant/case characteristics

Case reference		Asylum Case identification number
		Applicant's unique identifier
		Date of Registration (Lodging)
Demographics		Citizenship
		Language
		Age group
		Gender
Type of case (Procedure and prioritization)		Type of procedure
		Prioritization case
		Reasons for Prioritization Case
		Subsequent application
Interviews	Initial schedule	Interview scheduled date (1)
		Type of Interview
		Interview scheduled outcome (1)
	End of interview	Interview scheduled date (2)
		Type of Interview
		Interview scheduled outcome (2)
	Special procedural guarantees	Special procedural guarantees during the interview
Legal Support		Legal Support during interview
Concluding remarks		Concluding remarks date
		Concluding remarks outcome
		Identified special procedural guarantees included in opinion
Official decision by [<i>national authority</i>]		Official decision date
		Official decision outcome
Interview duration	Interview 1	Duration of Slot
		Starting Time (HH:MM)
		Ending Time (HH:MM)
	Interview 2	Duration of Slot
		Starting Time (HH:MM)
		Ending Time (HH:MM)
Case officers and Location		case officer 1 (code)
		case officer 2 (code)
		Location of Asylum Team
		Location of case officer (Interview)

Source: Internal operational monitoring

ANNEX 4: Feasibility checklist

To support a contractor implementing a future trial, the table below summarises, for the six criteria used to assess feasibility in this Study, a checklist of questions and issues to consider, in the context of a specific Member State in which a trial is to be implemented.

Criteria	Relevant questions and steps needed to determine feasibility and inform RCT design
Is it feasible to define the 'intervention' and implement it with fidelity in operational contexts?	<ul style="list-style-type: none"> Is the trial only including interviews for eligibility (not admissibility) as recommended in this Study? If not, clearly state which types of interviews are to be included. Are the definitions of remote and face-to-face interviews suggested in this Study applicable and acceptable in the Member State? If not, amendments should precisely define the modalities in a way that would allow replication. Develop programme theories for remote and face-to-face interviews in the Member State, including all the applicable inputs, activities, outputs, outcomes (corresponding to the process maps). This should also include the barriers and facilitators to implementation. List the role of all stakeholders involved in the applicant journey in the Member State. Describe their work in enough detail to allow replication. Develop a process map of the application process in the Member State – a more detailed version of Figure 3. If processes differ within the Member State, then separate process maps should be developed. Overall, ensure there is a clear statement about how remote and face-to-face interviews should be implemented in the Member State – in a way which is coherent with the process maps, programme theory and relevant national and EU legal standards. This should include clear indicators for fidelity of implementation.
Are there feasible options for randomisation?	<p>Randomising case officers:</p> <ul style="list-style-type: none"> Identify the number of case officers in post. Estimate likely numbers of case officers who would leave their post during the trial. Define inclusion/exclusion criteria (for example, whether to include only case officers with a certain level of experience in the trial). Identify the location of case officers within the Member State, and the extent to which they move between locations. Identify/procure an app/spreadsheet to use for the random allocation. Decide if the sample will be stratified. This decision could be based on the location of case officers or based on the findings of analysis of the statistical properties of the QAT (for example, if this showed that case officer experience was correlated with interview quality). Decide the randomisation ratio, using information about the average time needed for each modality, the available facilities, etc. <p>Randomising applicants:</p> <ul style="list-style-type: none"> Define inclusion/exclusion criteria (for example, excluding minors under 15, vulnerable applicants, etc.). Identify the point at which assessment of eligibility and randomisation will occur in the national process and who will be responsible for this. Define clear procedures for eligibility assessment and randomisation. Provide training to the staff who will undertake eligibility screening and randomisation. Decide if the sample is to be stratified (as with case officers, stratification might be used based on the physical location of applicants and/or analysis into how applicant characteristics are correlated with QAT score). Decide the randomisation ratio (possibly one for the case officer level and one for the applicant level). Use a pilot phase to test randomisation processes and amend if needed.
Is it feasible to define a clear primary and secondary research question and to collect	<ul style="list-style-type: none"> Does the primary, non-inferiority question (and the secondary question) recommended in this Study answer the question of interest to the national authority of the Member State in question? If not, why not, and what is missing?

data/to answer those questions?	<ul style="list-style-type: none"> • Is the EUAA QAT already used in the Member State? If not, could it be used for the purposes of the trial? Are there enough trained assessors? • If the QAT is being used, undertake an investigation of its statistical properties – if possible, using historic data from that Member State. • Ensure that the QAT app (if used) is able to output the data in the format needed for the trial – including a case identifier needed to link to other data sets. • Decide how to operationalise the output of the QAT (the categorical high/medium/low or the continuous numeric quality score). • Ensure that the quality assessors have access to the QAT app, are fully trained, and speak the language used in the interview transcripts. • Identify the data owner of the transcripts, whether they are in paper or electronic form, and ensure that quality assessors can gain access to the transcripts for the purposes of the trial – including preparing data sharing agreements if needed. • Decide whether it is feasible to redact transcripts so that the assessor is blind to the modality. • Decide if interviews should be quality scored by two assessors, as a check on consistency, and if so, what proportion of transcripts should be double coded. • Identify the owners of the other kinds of data needed for the trial (applicant and case officer characteristics), identify how these can be shared with the contractor appointed to conduct the trial, how they can be linked, including preparing data sharing agreements if needed. • Use the pilot period of the trial to practice and refine the linkage process.
Is the number of interviews conducted likely to generate a sufficient sample size within a reasonable timeframe?	<ul style="list-style-type: none"> • Work with the national authority and EUAA to agree an inferiority threshold. • Determine whether the outcome measure is a continuous variable (i.e., a numeric quality score), a categorical variable (i.e., high, low, or medium quality) or a binary variable (i.e., acceptable quality/unacceptable quality). • Understand the statistical properties of the QAT (mean scores and standard deviation, how interview quality varies by applicant and case officer characteristics). • Determine the inter-cluster correlation (comparing the variance within clusters, i.e., case officers, with the variance between clusters) • Identify the number of case officers who could participate in a trial, in a Member State. • Identify the number of interviews conducted per week per case officer, in the Member State. • Undertake power calculations, and work with the national authority and the EUAA to ensure a shared understanding about how long the trial would need to be implemented to reach the required sample size.
Is it feasible to implement a RCT ethically?	<ul style="list-style-type: none"> • Are remote and face-to-face interviews already used in the Member State, and if so, how frequently are these different modalities used? What are the views of the national authority about the possible risks to applicants from these different modalities? • Identify the ethical review body from which ethical approval for the trial would be needed. If possible, seek prior, informal advice from this awarding body. • Prepare a first draft of an ethical protocol, listing in detail the possible risks, harm (or benefits) to all those involved, along with mitigation measures. Share this draft with the national authority, the EUAA, and other relevant stakeholders to ensure a clear and shared understanding of the risks and mitigation measures. • Discuss the issue of whether or not to require applicants' informed consent. If consent is required, identify the point at which informed consent will be administered in the national process and who will be responsible for this. • Define clear procedures for administering informed consent and provide training to the staff who will undertake eligibility screening and randomisation. • Define a plan for monitoring adverse events throughout the trial, as well as conditions under which the trial should be discontinued. • Prepare a detailed data protection plan. Specifying all the information flows, any required data sharing agreements, how the confidentiality of applicants will be secured, etc.
Is it feasible to implement a RCT in operational, procedural and political contexts?	<ul style="list-style-type: none"> • Ensure fully informed support for the trial from that national authority – within relevant administrative departments, and possible at the political level. • Agree on a communications plan. • Identify reputational and other risks, and how to mitigate them.

