PREGNANCY- AND CHILDBIRTH-RELATED MEDIA USE TO SUPPORT MATERNAL MENTAL HEALTH

DATA SUMMARY

Will you re-use any existing data and what will you re-use it for?

No, we will collect new data.

What types and formats of data will the project generate or re-use?

study 1: interview data
Participants will be interviewed in person or over zoom. The interviews will be audio-recorded, and which will be transcribed verbatim by the PI (Femke Geusens). After transcribing the audio files, she will re-listen to the interviews to ensure that the transcriptions are complete and accurate. These transcriptions will be used for data analysis. The files will be named the pseudonym of the participant + the data the interview was conducted (e.g., Femke_20230109). Each participant's transcript will be stored in a separate file. Audiofiles will be deleted after transcription, so only the pseudonymized transcripts remain in existence.
study 2: panel data
The longitudinal panel study will be embedded in the running Mom2B study. These data are survey data with measures at four time points.

What is the purpose of the data generation or re-use and its relation to the objectives of the project?

study 1: interview data
The interview data will be used to answer three research questions:

- How do women with pregnancy-related anxiety make sense of and experience the different types of content they encounter on YouTube?
- How do they deal with potential unreliability of this content?
- Which specific content do they seek and avoid, depending on which needs they are trying to fulfill?

These research questions correspond to the first objective of the project: To understand how pregnancy- and childbirth-related content on YouTube can fulfill different needs for women with pregnancy-related anxiety.
study 2: panel data
The longitudinal panel data will be used to answer three research questions:

- Which underlying mechanisms explain how exposure to pregnancy- and childbirth-related YouTube content affects pregnancy-related anxiety?
- Do different types of content result in different effects via different processes?
- To what extend is the association between pregnancy-related anxiety and seeking pregnancy- and childbirth-related content on YouTube transactional

These research questions correspond to the first objective of the project: To elucidate the underlying processes of the association between exposure to pregnancy- and childbirth-related content and pregnancy-related anxiety.

What is the expected size of the data that you intend to generate or re-use?

study 1: interview data
We expect to interview between 20 and 50 participants, depending on when theoretical saturation is reached. Each interview is expected to last between 30 and 60 minutes. MP3 audio files are typically up to 150 MB per hour, meaning that the total size of our audio files should be no more than 10 GB. Transcripts of a one-hour interview are typically around 100-150 KB, meaning that the total size of the transcripts should be less than 10 MB.
study 2: panel data

We will recruit approximately 1000 participants, using measurements at 4 time points. A dataset of this size is approximately 15 MB as a .sav file (usable in SPSS).

What is the origin/provenance of the data, either generated or re-used?

All data is generated for this study

To whom might your data be useful ('data utility'), outside your project?

Data might be useful to other researchers studying maternal mental health, as well as healthcare providers and intervention specialists working around maternal mental health.

FAIR DATA

2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?

Pseudonymized or anonymized data linked to published articles will be made available on OSF, as well as a description of the data analysis (e.g., syntax and output files for quantitative data analysis). The OSF project page will be identified by a persistent identifier. No identifiable data will be shared or stored on OSF.

2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Yes, we will follow the DDI (Data Documentation Initiative) standard, which is an international standard for describing data from the social, behavioral, and economic sciences.

2.1. Making data findable, including provisions for metadata: Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?

yes

2.1. Making data findable, including provisions for metadata: Will metadata be offered in such a way that it can be harvested and indexed?

No

2.2. Making data accessible - Repository: Will the data be deposited in a trusted repository?

Yes, pseudonymized or anonymized data linked to published articles will be made available on OSF, as well as a description of the data analysis (e.g., syntax and output files for quantitative data analysis). The OSF project page will be identified by a persistent identifier. No identifiable data will be shared or stored on OSF.

2.2. Making data accessible - Repository: Have you explored appropriate arrangements with the identified repository where your data will be deposited?

Everyone can create a project page on OSF where they can deposit their data. I already have an account on OSF and have used them in the past.

2.2. Making data accessible - Repository: Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?

On OSF, you can choose whether the data is assigned an identifier.

2.2. Making data accessible - Data:
Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

Only pseudonymized or anonymized data will be made openly available. While the project is running, data will only be made available in smaller datasets containing only the variables used in specific papers, so as to give the research team the benefit of being the first to use their own data as a whole to test their hypotheses and explore their research questions. Once the project has finished and the hypotheses and research questions are answered, the full (pseudonymized or anonymized) dataset will be made available upon reasonable request.

2.2. Making data accessible - Data:
If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Smaller datasets containing only the variables used in specific papers will be made available as soon as the paper is submitted for review.

2.2. Making data accessible - Data:
Will the data be accessible through a free and standardized access protocol?

Yes, smaller datasets containing only the variables used in specific articles will be made available for free on OSF. The full dataset will be made available upon reasonable request, after a research plan has been approved by the researchers.

2.2. Making data accessible - Data:
If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?

A data sharing agreement will be drawn up and signed by both parties when providing access to the dataset. The smaller datasets made available during the project will include a statement that the data is only allowed to be used for scientific purposes.

2.2. Making data accessible - Data:

How will the identity of the person accessing the data be ascertained?

When the data are deposited in an open repository such as OSF, the identity of the person accessing the data cannot be ascertained. This is why no identifying data will be made available in these files. When access to the full dataset is requested, researchers will have to propose an research plan for which they want to use the data. This will allow us to

ascertain that the individual requesting access to the data is indeed a researcher. The anonymized dataset will then be shared with them using an individual link over a secure server.

2.2. Making data accessible - Data:

Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?

No, the PI can evaluate and approve access requests since the data will all be anonymized.

2.2. Making data accessible - Metadata:

Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

Yes

2.2. Making data accessible - Metadata:

How long will the data remain available and findable? Will metadata be guaranteed to remain available after data is no longer available?

Data will remain available and findable for 10 years. OSF guarantees availability of the metadata for 50 years.

2.2. Making data accessible - Metadata:
Will documentation or reference about any software be needed to access or read the data be included? Will it be possible to include the relevant software (e.g. in open source code)?

SPSS will be needed to read the datasets, syntaxes and outputs.

2.3. Making data interoperable:
What data and metadata vocabularies, standards, formats or methodologies will you follow to make your data interoperable to allow data exchange and re-use within and across disciplines? Will you follow community-endorsed interoperability best practices? Which ones?

A measurement binder with the original citation of the validated scales, as well as the specific items and answer options will be provided.

2.3. Making data interoperable:
In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies? Will you openly publish the generated ontologies or vocabularies to allow reusing, refining or extending them?

not applicable

2.3. Making data interoperable:
Will your data include qualified references[1] to other data (e.g. other data from your project, or datasets from previous research)?

[1]*A qualified reference is a cross-reference that explains its intent. For example, X is regulator of Y is a much more qualified reference than X is associated with Y, or X see also Y. The goal therefore is to create as many meaningful links as possible between (meta)data resources to enrich the contextual knowledge about the data. (Source:* https://www.go-fair.org/fair-principles/i3-metadata-include-qualified-references-metadata/***)***

yes, if needed

2.4. Increase data re-use:
How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?

These files will be uploaded alongside the data on the OSF project page. We will provide a measurement binder explaining the variables and answer options, describe the data cleaning process, and share the syntaxes used for data analysis.

2.4. Increase data re-use:
Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?

Data will be made available upon reasonable request.

2.4. Increase data re-use:
Will the data produced in the project be useable by third parties, in particular after the end of the project?

yes, and data will be made available upon reasonable request.

2.4. Increase data re-use:

Will the provenance of the data be thoroughly documented using the appropriate standards?

yes

2.4. Increase data re-use:

Describe all relevant data quality assurance processes.

We will have frequent “data meetings” where we try to visualize data, see number of missing, ranges (if realistic), and then proceed to data cleaning/management.

2.4. Increase data re-use:

Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

The only other research output that can be reused is the interview guide which will be used for the semi-structured interviews. This interview guide will be shared on OSF.

OTHER RESEARCH OUTPUTS

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

The only other research output that can be reused is the interview guide which will be used for the semi-structured interviews. This interview guide will be shared on OSF.

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

The only other research output that can be reused is the interview guide which will be used for the semi-structured interviews. This interview guide will be shared on OSF, making it findable, accessible, interoperable and reusable.

ALLOCATION OF RESOURCES

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?

considering that OSF is a free service, we do not foresee any costs for making the data FAIR.

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

not applicable

Who will be responsible for data management in your project?

Femke Geusens (PI)
Alkistis Skalkidou (advisor)

How will long term preservation be ensured? Discuss the necessary resources to accomplish this (costs and potential value, who decides and how, what data will be kept and for how long)?

We use BIANCA and VESTA for long term storage. Both are encrypted, you need authentication procedures to reach them, few have access and access is revised every year, data are always pseudo-anonymized and code keys are kept in totally different servers, data are backed-up regularly.
The smaller anonymous datasets made available on OSF. OSF is a trusted repositorywhich assigns every project, component, and file a short url that is globally unique and whose persistence is guaranteed by a data preservation fund currently sufficient to provide 50+ years of public access.

DATA SECURITY

What provisions are or will be in place for data security (including data recovery as well as secure storage/archiving and transfer of sensitive data)?

We use BIANCA, UPPMAX and VESTA to securely store our data. All three are encrypted, you need authentication procedures to reach them, few have access and access is revised every year, data are always pseudo-anonymized and code keys are kept in totally different servers, data are backed-up regularly.

ETHICS

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Identifying data, such as the raw data including email addresses or audiofiles will never be shared. Only pseudonymized or anonymized data without reference to identifying data will be shared.

Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

yes

OTHER ISSUES

Do you, or will you, make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones (please list and briefly describe them)?

/