Consensus Recommendations to Establish Reporting Standards in fMRI of Migraine

A Delphi Study

Severin Schramm,[1] Corinna Börner-Schröder,[1,2,3] Miriam Reichert,[1] Constanze Ramschütz,[1] Xiao Michelle Androulakis,[4] Messoud Ashina,[5,6] Gianluca Coppola,[7] Brett Cucchiara,[8] Zhao Dong,[9] Xiaoxia Du,[10] Laura H. Fischer-Schulte,[11] Peter J. Goadsby,[12] Rune Häckert Christensen,[5,6,13,14] Luke A. Henderson,[15] Anders Hougaard,[6,16] Jian-Ren Liu,[17] Gabriella Juhasz,[18,19] Nazia Karsan,[12] Jian Kong,[20] Jeungchan Lee,[21,22] Mi Ji Lee,[23] Clas Linnman,[24] Vani Mathur,[25] Arne May,[26] Jan Mehnert,[26] Eric Moulton,[27,28] David M. Niddam,[29] Jean Schoenen,[30] David A. Seminowicz,[31] Anne Stankewitz,[1] Yiheng Tu,[32] Dániel Veréb,[33,34] Tao Yin,[35] Claus Zimmer,[1,36] Florian Heinen,[2] Thomas Baum,[1] Michaela V. Bonfert,[2,3] and Nico Sollmann[1,36,37]

Correspondence
Dr. Sollmann
nico.sollmann@tum.de

Neurology® 2025;104:e210235. doi:10.1212/WNL.0000000000210235

Abstract

Background and Objectives

Migraine is a multifaceted primary headache disorder. In neuroimaging of migraine, fMRI has been used to elucidate pathophysiology or monitor treatment effects. The current literature, however, is highly heterogeneous regarding reported variables and methodologies. This begets a lack of comparability and complicates synthesis of results across studies. We developed a framework for standardized reporting of fMRI studies in migraine.

Methods

Experts on fMRI in migraine were identified from the literature and subjected to structured questionnaires in 2 iterations of 3 rounds according to the DELPHI method. A total of 157 statements across 17 reporting domains were rated on 5-point Likert scales (strong support to strong opposition). The first iteration covered demographic data, migraine-specific factors, medication, scan timing, healthy controls (HCs), participant sampling/recruiting, standardized forms, study preregistration, region of interest (ROI) analyses, validation data sets, data sharing, preprocessing documentation, and analysis software. The second iteration of the questionnaire covered scanner-related factors, sequence-related factors, physiology monitoring, and stimulation-related factors. Items showing strong consensus/consensus (≥90%/≥75% of participants indicating scores 4 or 5) were included as standard reporting items.

Results

All 3 rounds of the first/second iteration were completed by 29 and 26 researchers (age 46 ± 11 years; 38% female/age 46 ± 12 years; 44% female) from 23 and 21 institutions. Across both iterations, strong consensus and consensus was achieved for 34 (3 scanner-related factors,

[1]Department of Diagnostic and Interventional Neuroradiology, School of Medicine, Klinikum rechts der Isar, TUM School of Medicine and Health, Technical University of Munich, Germany; [2]Department of Pediatric Neurology and Developmental Medicine, Dr. von Hauner Children's Hospital, LMU Hospital, Munich, Germany; [3]LMU Center for Children with Medical Complexity - iSPZ Hauner, Ludwig Maximilian University, Munich, Germany; [4]Neurology Division, Columbia VA Health Care System, SC; [5]Department of Neurology, Danish Headache Center, Rigshospitalet Glostrup, Denmark; [6]Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark; [7]Department of Medico-Surgical Sciences and Biotechnologies, Sapienza University of Rome Polo Pontino ICOT, Latina, Italy; [8]Department of Neurology, University of Pennsylvania, Philadelphia; [9]Department of Neurology, Chinese PLA General Hospital, Beijing, China; [10]School of Psychology, Shanghai University of Sport, China; [11]Clinic and Policlinic of Psychiatry, University Medical Center Hamburg-Eppendorf, Hamburg, Germany; [12]National Institute for Health Research - Wellcome Trust King's Clinical Research Facility, King's College London, United Kingdom; [13]Department of Anesthesia, Critical Care and Pain Medicine, Beth Israel Deaconess Medical Center, Boston, MA; [14]Harvard Medical School, Boston, MA; [15]School of Medical Sciences (Neuroscience), Brain and Mind Centre, The University of Sydney, New South Wales, Australia; [16]Department of Neurology, Copenhagen University Hospital - Herlev & Gentofte, Denmark; [17]Department of Neurology, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, China; [18]Department of Pharmacodynamics, Semmelweis University, Budapest, Hungary; [19]NAP3.0 Neuropsychopharmacology Research Group, Semmelweis University, Budapest, Hungary; [20]Department of Psychiatry, Massachusetts General Hospital, Boston; [21]Department of Radiology, Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston; [22]Discovery Center for Recovery from Chronic Pain, Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Boston, MA; [23]Department of Neurology, Seoul National University Hospital, South Korea; [24]Department of Physical Medicine and Rehabilitation, Spaulding Rehabilitation Hospital, Boston, MA; [25]Department of Psychological and Brain Sciences, Texas A&M University, College Station; [26]Department of Systems Neuroscience, University Medical Center Hamburg-Eppendorf, Germany; [27]Department of Anesthesiology, Critical Care and Pain Medicine, Boston Children's Hospital, Boston, MA; [28]Department of Ophthalmology, Boston Children's Hospital, Harvard Medical School, Boston, MA; [29]Brain Research Center, National Yang Ming Chiao Tung University, Taipei, Taiwan; [30]Headache Research Unit, University of Liège, Belgium; [31]Department of Medical Biophysics, Schulich School of Medicine & Dentistry, University of Western Ontario, London, Canada; [32]Institute of Psychology, Chinese Academy of Sciences, Beijing, China; [33]Department of Clinical Neuroscience, Karolinska Institutet, Stockholm, Sweden; [34]Department of Radiology, Albert Szent-Györgyi Health Centre, Szeged, Hungary; [35]Acupuncture and Tuina School/Acupuncture-Brain Science Research Center, Chengdu University of Traditional Chinese Medicine, China; [36]TUM-Neuroimaging Center, Klinikum rechts der Isar, Technical University of Munich, Germany; and [37]Department of Diagnostic and Interventional Radiology, University Hospital Ulm, Germany.

Glossary

BOLD = blood oxygen level dependent; **HC** = healthy control; ***ICHD*** = *International Classification of Headache Disorders*; **ROI** = region of interest.

9 sequence-related factors, 1 stimulation-related factor, 2 demographic factors, 7 migraine-specific factors, 2 medication-factors, 2 scan timing factors, 4 HC factors, 1 preregistration factor, 1 analysis software factor, and 2 ROI analyses factors) and 33 (1 scanner-related factors, 4 sequence related factors, 1 factor related to physiology monitoring, 1 stimulation-related factor, 3 demographic factors, 6 migraine-specific factors, 4 medication factors, 3 HC factors, 2 sampling factors, 1 standardized form, 1 preregistration factor, 1 data sharing factor, 2 analysis software factors, and 3 ROI analyses factors) items, respectively. From these, a checklist covering 63 items from 14 reporting domains was created.

Discussion

We present an expert-based framework for reporting standards in fMRI studies of migraine, which can be used for future studies to homogenize cohort characterization, fMRI acquisitions, and analysis protocols.

Introduction

Migraine is a primary headache disorder codified in the *International Classification of Headache Disorders* (*ICHD*)[1] affecting approximately 15% of the population.[2-4] The disease causes extensive disability, corresponding to 4.9% of global years lived with disability and 1.7% of disability-adjusted life years in 2019.[2,4]

fMRI exploits neurovascular coupling to measure neuronal activity indirectly using the blood oxygen level dependent (BOLD) signal.[5,6] Although widely used in scientific and clinical investigations, methodological concerns such as a lack of reproduction of findings, underpowered studies, and lack of data access cause intense debates within the field.[7,8] Overall, this leads to a situation in which confidence in the validity of many findings is considerably diminished.

In this context, the large number of clinically relevant dimensions inherent in migraine (e.g., migraine subtype, attack frequency, pain intensity, presence of medication, and aura characteristics) poses particular challenges, with documentation standards regarding many of these variables being inconsistent between studies.[9,10] Additional inconsistencies are introduced through methodological heterogeneities such as varying sample sizes, inconsistent sequence parameters, and different approaches to statistical analyses.

To address this issue, we aimed to develop a framework for reporting of fMRI studies in migraine to improve the quality of the future literature landscape by enabling researchers to consider and document relevant factors of interest. Specifically, we used the DELPHI approach to achieve consensus-based recommendations to establish reporting standards for fMRI in migraine.

Methods

To generate reporting standards in fMRI of migraine, we followed the DELPHI method.[11,12] Essential components of the DELPHI method include the assembly of an expert group, the conduction of multiple iterations of questionnaires to collect opinions on the issues to be appraised, and a feedback mechanism to inform the questioned individuals about the results of previous iterations, while keeping the identity behind individual replies anonymous.[11]

Standard Protocol Approvals, Registrations, and Participant Consents

All researchers were informed about the purpose of their involvement and gave written consent to take part in the data collection.

Selection of Participants

Participants were selected through the literature pool from a recent systematic review of fMRI in migraine.[10] The review encompassed 114 articles published between April 2014 and January 2021.[10]

The first and last authors (all authors in case of shared first or senior authorship) of each article included in the review were extracted, and corresponding email addresses were acquired by web research. In this way, we ensured that all participants were experts who had published peer-reviewed work related to fMRI investigations of migraine. The primary data collectors and initiators (S.S., C.B.S., C.R., M.R., C.Z., F.H., T.B., M.B., and N.S.) did not take part in the questionnaires.

DELPHI Questionnaire

We followed established guidelines for consensus generation by the DELPHI method.[11] In total, we conducted 2 iterations of 3 rounds; that is, we consecutively prepared and shared 2 distinct questionnaires covering 3 rounds each, with results from each round informing the subsequent questionnaire. The first iteration focused primarily on demographic, clinical, and study design aspects while the second iteration focused primarily on scanner-related and imaging-related aspects. Both questionnaires were distributed through Google Forms.

Participants were offered a co-author position on this article for taking part in all 3 questionnaire rounds. The period for consecutive completion of all 3 questionnaires ranged from July 7, 2022, to June 7, 2023 (iteration 1), and from April 24 to July 12, 2024 (iteration 2). A complete list of all questions can be found in eTables 1 and 2.

Experts rated their agreement/disagreement regarding specific items on 5-point Likert scales, with scores for strongly agree (5), rather agree (4), neutral (3), rather disagree (2), and strongly disagree (1). In total, 157 items were rated this way.

We defined "strong consensus" for items when ≥90% of participants indicated Likert scores 4 or 5 (support or strong support) and "consensus" for items when ≥75% of participants indicated Likert scores 4 or 5 (support or strong support).

Questionnaire of Round 1

The aim of round 1 regarding both iterations was a qualitative primary survey of opinions on a variety of potential topics of interest, both to generate items for consensus quantification during round 2 and to get a preliminary sense of opinion distributions.

Specifically, we used 50 open-ended questions on the general role of fMRI in migraine research, a number of study design considerations (determination of sample size, value of study preregistration, data preprocessing pipelines, region of interest [ROI] selection, software selection, statistical analysis, and validation of test results), data availability, general demographic variables, migraine-specific variables, scan timing in relation to the migraine cycle, medication, and standardized forms (eTable 1).

For round 1 of iteration 2, more methodological and technical domains of interest regarding fMRI were covered. This round included 11 open-ended questions on scanner properties, sequence design, and physiology monitoring (eTable 2).

The replies were reviewed by S.S. and N.S. Items rated according to the Likert scales were generated for each potential item identifiable from the replies and specific enough to fMRI of migraine to be considered relevant for a reporting consensus.

Questionnaire of Round 2

The aim of round 2 for both iterations was threefold, providing feedback from round 1 to the participants to inform them on the previous distribution of opinions, generating opinion distribution data on items that emerged from round 1, and generating additional items for round 3 based on feedback from rounds 1 and 2.

We submitted 102 statements to be rated on the 5-point Likert scales, indicating the level of agreement or disagreement for a respective item ranging from strong agreement to strong disagreement. In addition, we submitted 4 open-ended questions regarding the topic of healthy controls (HCs), which emerged as a topic of interest in round 1 of iteration 1, and 1 open-ended question regarding the definition of scan timing in relation to migraine attacks (eTable 1).

Likewise, round 2 for iteration 2 was conducted to collect quantifiable consensus data with a focus on technical aspects. The round covered 33 scaled items on scanner-related factors, sequence-related factors, and physiology monitoring. It also contained an open-ended question related to stimulation-based fMRI paradigms (eTable 2).

Questionnaire of Round 3

The aim of round 3 of both iterations was to complete the DELPHI process by getting opinion distribution data on any residual items that had emerged from the previous 2 rounds. For iteration 1, we submitted 17 statements regarding characterization and selection of HCs, fMRI preprocessing pipelines, and standardized questionnaires to be rated on a Likert scale similar to round 2 (eTable 1).

For iteration 2, a second third-level questionnaire was administered with a focus on technical fMRI aspects. This round consisted of 4 scaled items related to stimulation-based fMRI paradigms and 1 scanner-related factor (eTable 2).

Reporting Domains

Based on the collected replies, we assembled 17 distinct reporting domains encompassing all 157 ranked items. In total, we collected information on the domains of demographic data (24 items), migraine-specific factors (32 items), medication (10 items), scan timing (3 items), HCs (12 items), participant sampling/recruiting (3 items), standardized forms (11 items), study preregistration (5 items), ROI analyses (5 items), validation data sets (2 items), data sharing (2 items), preprocessing documentation (7 items), and analysis software (3 items), as well as scanner-related factors (5 items), sequence-related factors (18 items), factors related to physiology monitoring (11 items), and stimulation-related factors (4 items).

Statistical Analysis

Replies from each round were exported from Google Forms to a Microsoft Excel file (version 16.61.1, 2022; Microsoft Corp., Redmond, WA) for further analysis. All Likert-scaled items were analyzed regarding the distribution of opinions, and items that reached our previously defined levels of consensus were extracted for specific reporting of descriptive statistics (i.e., description of opinion distribution frequencies) by S.S. and N.S.

Data Availability

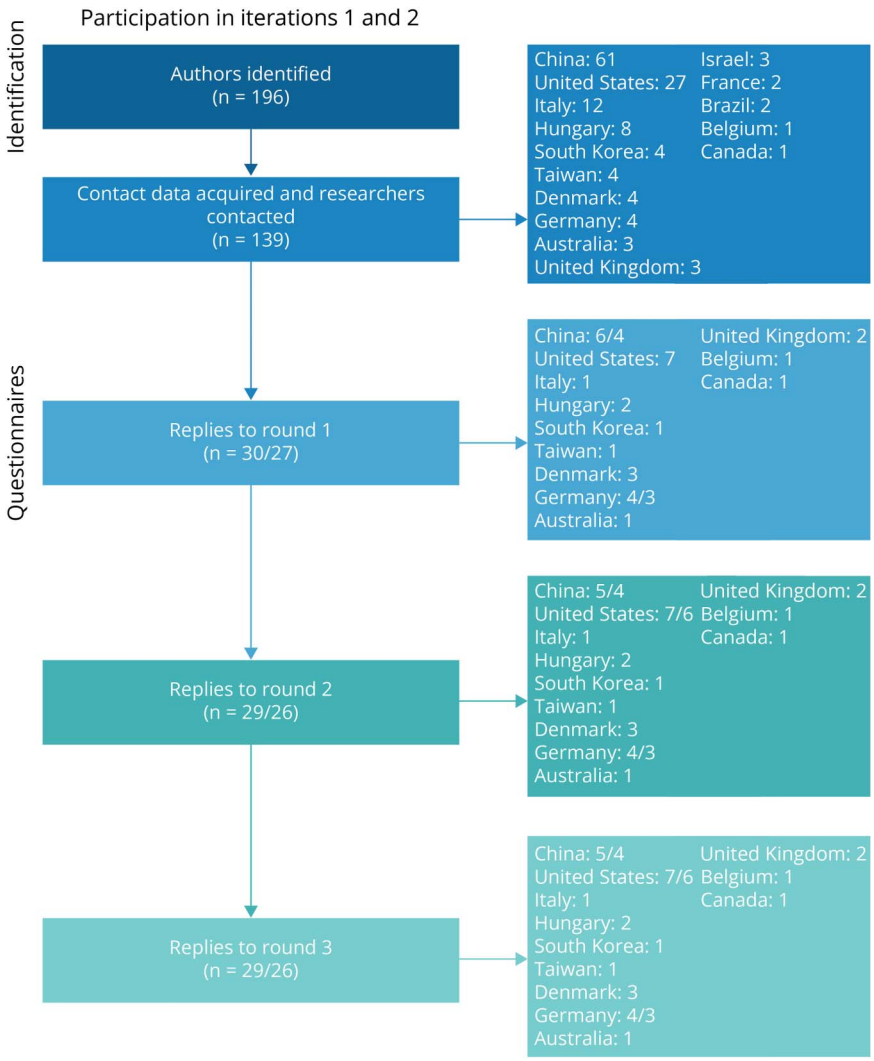
The entirety of our data is, in anonymized fashion, available in eTables 1 and 2.

Results

Summary of Replies

From our literature base, we extracted 196 potential experts. The Figure demonstrates the flow of replies including

Figure Participant Identification and Reply Flow



The number and national affiliations of contacted experts at different stages of the DELPHI process during both iteration 1 and iteration 2 are demonstrated. In cases where the number of experts differs between iteration 1 and 2, the respective numbers are indicated by "/".

national affiliation for both iterations. Email contacts were identified for 139 of these experts (71%).

After submitting our invite for the questionnaire of round 1 of iteration 1, a total number of 30 contacts replied (22%). For the questionnaire of round 2, 29 contacts replied (97%). For the questionnaire of round 3, 29 contacts replied (100%). The average participant age in iteration 1 was 46 ± 11 years, with a female/male ratio of 8/21. Participants were spread over 23 institutions from 12 nations and 4 continents (Asia [28%], Europe [45%], North America [24%], and Australia [3%]).

For iteration 2, 26 experts took part in all 3 rounds. The average age was 46 ± 12 years, with a female/male ratio of 8/18 and a spread over 21 institutions from 12 nations and 4 continents (Asia [23%], Europe [46%], North America [27%], and Australia [4%]).

Information on participants is presented in eAppendix 1. It is important to note that for selected items, some researchers abstained from voting; the respective information is presented in eTables 1 and 2 and eAppendix 2.

Reporting Items

Overall, strong consensus was achieved for 34 items (22%) and consensus for 33 items (21%) of 157 total items (all rounds of iteration 1 and iteration 2 together). A simple majority (>50% of Likert scores 4 or 5) was present for 44 items (28%).

Demographic Factors

Regarding general demographic factors, we collected opinion distribution data on 24 total items (eTable 1). Strong consensus was achieved for 2 and consensus for 3 items (Table 1).

Migraine-Specific Factors

For disease-specific factors, we collected opinions on 32 items (eTable 1). We achieved strong consensus for 7 items and consensus for 6 items (Table 1).

Table 1 Overview of Reporting Items With Strong Consensus or Consensus

	Items	
Domains	**Strong consensus**	**Consensus**
Demographics	Age, sex	Depression scores, anxiety scores, menstrual cycle
Migraine	*ICHD*-based diagnosis, neurologic and psychiatric comorbidities, migraine frequency, disease duration, headache frequency, time from last headache to scan, pain comorbidities	Migraine subtype, aura characteristics, symptom laterality, photophobia, mean pain intensity, phonophobia
Medication	Types of prophylactic medication, acute medication before scanning	Prophylactic medication before scanning, types of acute medication, baseline frequency of acute medication, baseline frequency of prophylactic medication
Healthy controls	Matching by age, no headache disorders, no pain disorders, matching by sex	No neurologic diseases, no psychiatric diseases, matching by as many variables as possible
Scan timing	Nomenclature according to *ICHD* definitions of prodrome, postdrome, and interictal intervals, reporting of scan timing in migraine cycle	—
Sampling	—	Recruitment procedures, limitations due to specific recruitment
Standardized forms	—	Symptom documentation using diagnostic headache diary
Study preregistration	More negative findings should be published	Exploratory non-preregistered analyses are valuable
ROI analyses	Exact coordinates and extent, strong foundations in underlying hypotheses	ROI placement informed by previous work, discuss impact of ROI selection process, atlas can facilitate replicability
Validation data sets	—	—
Data sharing	—	Data sharing should become the norm
Preprocessing	—	—
Analysis software	Justification if not open source	Open source preferred, code should be shared
Scanner	Data harmonization methods, static field strength, head coils	Scanner model
Sequence	Voxel size, repetition time, slice thickness, in-plane resolution, scanned anatomy, flip angle, number of slices, scan duration, image acquisition method	Echo time, acquisition matrix, number of volumes acquired, image orientation
Physiology monitoring	—	Movement minimization procedure
Stimulus	Essential stimulus characteristics	Stimulus presentation devices

Abbreviations: *ICHD* = *International Classification of Headache Disorders*; ROI = region of interest.
Table 1 yields an overview over all reporting items that have reached consensus or strong consensus. More detail is presented within the supplementary materials.

Medication

For the 10 items related to migraine medication (eTable 1), strong consensus was achieved for 2 items and consensus for 4 items (Table 1).

Scan Timing of fMRI

Regarding the reporting of scan timing in relation to migraine attacks (eTable 1), 3 items were submitted, of which 2 items achieved strong consensus (Table 1).

Because the items referenced prodrome and postdrome, we want to reiterate that the *ICHD* currently defines prodrome as a "symptomatic phase, lasting up to 48 hours, occurring before the onset of pain in migraine without aura or before the aura in migraine with aura. Among the common prodromal symptoms are fatigue, elated or depressed mood, unusual hunger, and cravings for certain foods."[1] The postdrome is defined as a "symptomatic phase, lasting up to 48 hours, following the resolution of pain in migraine attacks with or without aura. Among the common postdromal symptoms are fatigue, elated or depressed mood, and cognitive difficulties."[1]

Healthy Controls

In total, 12 items related to the selection of HCs were submitted for evaluation (eTable 1). Of those, 4 achieved strong consensus and 3 consensus (Table 1).

Sampling

We submitted 3 items for consideration that related to sampling procedures (eTable 1). Of these, 2 items achieved consensus (Table 1).

Standardized Forms

Regarding the potential inclusion of standardized forms or questionnaires in our reporting standards, we submitted 11 items for consideration (eTable 1). Consensus emerged for 1 item (Table 1).

Study Preregistration

For the topic of preregistration of study protocols, 5 items were included in the questionnaires (eTable 1). Strong consensus emerged on 1 item and consensus on 1 item (Table 1).

Data Sharing

Two items were submitted relating to the practice of data sharing (eTable 1). One of these items achieved consensus (Table 1).

Software for Analysis

Three of the submitted items referred to software used in fMRI analyses (eTable 1). 1 of these achieved strong consensus, and 2 achieved consensus (Table 1).

ROI Analyses

To generate consensus on principles for ROI-based analyses, we submitted 5 items (eTable 1). 2 of these reached strong consensus and 3 reached consensus (Table 1).

Preprocessing Documentation and Validation Data Sets

We collected opinion distributions on 7 items related to various standards of preprocessing documentation and 2 items related to validation data sets (eTable 1), of which no item reached consensus (Table 1).

Scanner-Related Factors

For reporting items related to the scanner model, 5 scaled items emerged from the open questions (eTable 2). Strong consensus was achieved for 3 and consensus for 1 factor (Table 1).

Sequence-Related Factors

In the context of sequence-related factors, the total number of items submitted was 18 (eTable 2). Strong consensus was achieved for 9 and consensus for 4 factors (Table 1).

Factors Related to Physiology Monitoring

Overall, 11 items related to physiology monitoring were submitted (eTable 2). Consensus was achieved only for 1 factor (Table 1).

Stimulation-Related Factors

Four items covered stimulation-related factors (eTable 2). Reporting of one item reached strong consensus, while another item reached consensus. (Table 1).

Discussion

In our pursuit to generate consensus recommendations on reporting standards in fMRI of migraine using the DELPHI method, we received replies from 29 experts (iteration 1) and 26 experts (iteration 2). Of 157 total items across iterations, we obtained strong consensus (≥90% of participants indicating Likert scores 4 or 5 [support or strong support]) on 34 items and consensus (≥75% of participants indicating Likert scores 4 or 5 [support or strong support]) for 33 items.

Correspondingly, we have designed a checklist including all items that achieved at least consensus and were suitable for a checklist. Four items were not included in the checklist because they dealt with recommendations not applicable to reporting within a specific study (e.g., "in general, analysis of fMRI data should be performed using open-source and validated software").

The list encompasses 63 items spanning 14 reporting domains (eAppendix 3). It is important to note that this list is necessary, but not sufficient. As always, reporting should allow for precise replication.

Migraine is subject to many clinical phenotypes.[13] This includes factors such as presence and character of aura, laterality, migraine intensity, frequency, disease cycle, and disease duration, among others.[1,10,13] These factors constitute degrees of freedom inherent in the disease but are oftentimes not considered in the reporting or design of imaging studies.

As noted in recent reviews, inconsistent reporting of these factors complicates the comparability of results across different studies.[9,10] In addition, the conduction of replicative studies to confirm previous findings is rendered increasingly difficult.

Reporting standards could help to ameliorate these circumstances by providing researchers with a guideline on factors that have been deemed essential by experts in the field. This would add comparability and structure to a field that could, given these changes, provide valuable insights into pathophysiologic processes underlying one of the most prevalent and disabling neurologic diseases worldwide. Our proposed reporting standards will not be able to address other sources of heterogeneity in the wider field of fMRI, such as insufficient sample sizes, heterogeneity in preprocessing pipelines, or increasing variety in possible analysis methods, collectively referred to as researcher degrees of freedom.[7] Naturally, the current minimum standards are not set in stone and should be revisited in time to be updated by new consensus.

While we were successful in finding consensus regarding many of the proposed items, we should also note that consensus was not found for other items. Notably, there was no consensus on testing initial results against unexamined validation data. Some experts supported this for reliability, an opinion shared by other researchers in the wider fMRI community,[7] while others disagreed, citing resource constraints. Despite recognizing the benefits, the drawbacks were deemed too significant for broad adoption. Solutions could include increased funding for validation or data sharing.

Furthermore, expert replies regarding the need for study preregistration were heterogeneous. While most experts indicated support for preregistration of study protocols because of potential benefits regarding transparency and the prevention of practices such as hypothesizing after results are known,[14] other experts noted limited benefits compared with the time investment. In addition, most researchers indicated that they had not performed detailed preregistration of their study protocols in their previous work. Regarding the reasons, researchers mostly referred to the practice not being widely spread during that time and the lack of requirements to do so.

Another problem related to study preregistration is the multitude of analysis methods available to examine BOLD-based imaging.[10] A large amount of analysis parameters can be freely chosen, which can complicate the interstudy comparability. Reflecting this, researchers indicated that documentation of preprocessing steps was a factor of concern, influencing analysis results. Nonetheless, although researchers referred to a diverse set of existing guidelines on the topic,[8,15-19] researchers did not establish consensus on which guidelines could serve as sufficient for recommendation in reporting standards. We believe that this reflects the need for standardization on technical aspects within the fMRI community as a whole.

Relatedly, while initially posing questions on statistical analyses, we realized that consensus would likely be difficult to achieve and might constrict experimental design. The plethora of potentially interesting analyses cannot easily be captured by rigid guidelines because the specific validity depends on the experiment and hypotheses. This again emphasizes the potential that preregistration of experiments and data analysis plans hold.[7]

An additional aspect touched on by our initial questionnaire was the topic of sample size determination. Opinions on the need for and validity of a priori power analyses in fMRI studies of migraine were heterogeneous. While some researchers indicated that sample size calculations would be strictly necessary, others emphasized the limiting effects of study costs and patient availability, rapidly changing standards in the field, and poor effect estimations. Again, clearer guidance from both the wider fMRI community and publishers may be required.

An interesting observation was made on validation data sets. During our open questionnaire, most researchers noted that testing of findings against previous validation data sets was generally desirable. Nonetheless, for the 2 items related to this topic, no consensus was achieved. Likely reasons for this dichotomy can be found in practical challenges, specifically identification of comparable data sets and sparsity of resources. Yet, some researchers noted that immediate self-validation may be less important if preregistration of the data analysis plan has been performed.

Finally, regarding demographic factors, consensus was limited to relatively few items. Previous publications have noted that in neuroimaging studies and medical science as a whole, factors potentially relevant to research questions such as race and ethnicity may be underreported.[20,21] Future reporting standards may thus expand on demographic factors.

Although we received responses from 29 experts in iteration 1 and 26 experts in iteration 2, this only corresponds to a response rate of 21% and 19% of all researchers contacted. Therefore, we cannot exclude the possibility of having missed input that may have contributed to more wide-ranging recommendations. One such limitation is visible in the high proportion of clinicians and relative paucity of technicians in our panel, which we can assume to be a source of bias (eAppendix 1). Considering the total number of 29 and 26 replies and the absence of other recommendations for fMRI in migraine, we are still confident that our reporting standards can provide an important contribution to the field.

Second, we only used 3 rounds of questioning per iteration. Additional rounds of questioning could potentially have further refined the present items and generated additional reporting items. Three rounds were chosen as a compromise between volume and precision of reporting items, as well as reasonable effort and time expense for the questioned experts.

Third, our reporting items are largely focused around clinical aspects, for example, exact symptom description, medication, or scan timing. Owing to controversies in the broader field of fMRI regarding technical aspects such as sample size, preprocessing, or optimal statistical analyses,[7] we have chosen to largely forgo recommendations within these areas because appropriate replies are likely to come from the wider field of fMRI rather than the subfield of migraine. Nonetheless, the impact of these methodological factors might be among the main factors underlying heterogeneous findings.[22,23] Previous work has outlined practical approaches to fMRI acquisition, data processing, and data analyses from a rather general perspective.[24]

In collecting and analyzing opinions of 29 and 26 experts (iterations 1 and 2) over 3 rounds of questionnaires using the DELPHI method, we have generated a checklist of 63 standardized reporting items spanning 14 domains related to fMRI investigations in migraine. Adoption of this checklist in future fMRI studies may improve interstudy comparability of results and replicability and could thereby contribute to a better understanding of migraine pathophysiology as investigated by fMRI.

Author Contributions

S. Schramm: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data. C. Börner-Schröder: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. M. Reichert: drafting/revision of the manuscript for content, including medical writing for content; study concept or design. C.

Ramschütz: drafting/revision of the manuscript for content, including medical writing for content. X.M. Androulakis: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. M. Ashina: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. G. Coppola: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. B. Cucchiara: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. Z. Dong: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. X. Du: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. L.H. Fischer-Schulte: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. P.J. Goadsby: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. R.H. Christensen: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. L.A. Henderson: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. A. Hougaard: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. J.-R. Liu: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. G. Juhasz: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. N. Karsan: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. J. Kong: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. J. Lee: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. M.J. Lee: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. C. Linnman: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. V. Mathur: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. A. May: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. J. Mehnert: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. E. Moulton: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. D.M. Niddam: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. J. Schoenen: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. D.A. Seminowicz: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. A. Stankewitz: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. Y. Tu: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. D. Veréb: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. T. Yin: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. C. Zimmer: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. F. Heinen: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. T. Baum: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. M.V. Bonfert: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data. N. Sollmann: drafting/revision of the manuscript for content, including medical writing for content; major role in the acquisition of data; study concept or design; analysis or interpretation of data.

Study Funding

No targeted funding reported.

Disclosure

S. Schramm, C. Börner-Schröder, M. Reichert, C. Ramschütz, and X.M. Androulakis report no disclosures relevant to the manuscript. M. Ashina acts as a consultant, speaker, or scientific advisor for AbbVie, Allergan, Amgen, Eli Lilly, Lundbeck, Novartis, and Teva; is a primary investigator for ongoing AbbVie/Allergan, Amgen, Eli Lilly, Lundbeck, Novartis, and Teva trials; has no ownership interest and does not own stocks of any pharmaceutical company; and serves as associate editor of *Cephalalgia*, associate editor of the *Journal of Headache and Pain*, and associate editor of *Brain*. G. Coppola, B. Cucchiara, Z. Dong, X. Du, and L.H. Fischer-Schulte report no disclosures relevant to the manuscript. P.J. Goadsby has received research grants from or served as a consultant to Abbott, Allergan, Almirall Prodesfarma, AstraZeneca, Bristol Myers Squibb, Elan, Glaxo SmithKline, Merck, Ortho-McNeil, Pfizer, Pharmacia, Pierre Fabre, and Vanguard. R.H. Christensen, L.A. Henderson, A. Hougaard, and J.-R. Liu report no disclosures relevant to the manuscript. G. Juhasz received funding from the Hungarian Brain Research Program (grant: 2017-1.2.1-NKP-2017-00002; NAP2022-I-4/2022) and from the Hungarian National Research, Development and Innovation Office (TKP2021-EGA-25 supported by the Ministry of Innovation and Technology of Hungary; 2019-2.1.7-ERA-NET-2020-00005 under the framework of ERA PerMed; K 143391). N. Karsan reports no disclosures relevant to the manuscript. J. Kong holds equity in startup companies (MNT, BTT) and a patent on applying neuromodulation but declares no conflict of interest. J. Lee,

M.J. Lee, C. Linnman, V. Mathur, A. May, J. Mehnert, E. Moulton, D.M. Niddam, J. Schoenen, D.A. Seminowicz, A. Stankewitz, Y. Tu, D. Veréb, T. Yin, and C. Zimmer report no disclosures relevant to the manuscript. F. Heinen received a grant "Innovationsfonds" of the joint federal committee of health insurance companies (GBA) for a nationwide study on an early multimodal intervention program for children with migraine. T. Baum reports no disclosures relevant to the manuscript. M.V. Bonfert is supported by the Bavarian Gender Equality Grant of the Free State of Bavaria, the German Migraine and Headache Society, and the ZNS-Hannelore-Kohl Stiftung. N. Sollmann is supported by the Dr.-Ing. Leonhard Lorenz Foundation, the German Migraine and Headache Society, and the Joachim Herz Foundation. Go to Neurology.org/N for full disclosures.

Publication History

Received by *Neurology* January 25, 2024. Accepted in final form December 11, 2024. Submitted and externally peer-reviewed. The editor was Associate Editor Rebecca Burch, MD.

References

1. Headache Classification Committee of the International Headache Society (IHS) The International Classification of Headache Disorders. *Cephalalgia*. 2018;38(1):1-211. doi:10.1177/0333102417738202
2. Stovner LJ, Nichols E, Steiner TJ, et al. Global, regional, and national burden of migraine and tension-type headache, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Neurol*. 2018;17(11):954-976. doi:10.1016/S1474-4422(18)30322-3
3. Fan L, Wu Y, Wei J, et al. Global, regional, and national time trends in incidence for migraine, from 1990 to 2019: an age-period-cohort analysis for the GBD 2019. *J Headache Pain*. 2023;24:79. doi:10.1186/s10194-023-01619-9
4. Vos T, Lim SS, Abbafati C, et al. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet*. 2020;396(10258):1204-1222. doi:10.1016/S0140-6736(20)30925-9
5. Logothetis NK. What we can do and what we cannot do with fMRI. *Nature*. 2008;453(7197):869-878. doi:10.1038/nature06976
6. Gauthier CJ, Fan AP. BOLD signal physiology: models and applications. *Neuroimage*. 2019;187:116-127. doi:10.1016/j.neuroimage.2018.03.018
7. Poldrack RA, Baker CI, Durnez J, et al. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat Rev Neurosci*. 2017;18(2):115-126. doi:10.1038/nrn.2016.167
8. Nichols TE, Das S, Eickhoff SB, et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nat Neurosci*. 2017;20(3):299-303. doi:10.1038/nn.4500
9. Skorobogatykh K, van Hoogstraten WS, Degan D, et al. Functional connectivity studies in migraine: what have we learned? *J Headache Pain*. 2019;20(1):108. doi:10.1186/s10194-019-1047-3
10. Schramm S, Börner C, Reichert M, et al. Functional magnetic resonance imaging in migraine: a systematic review. *Cephalalgia*. 2023;43(2):03331024221128278. doi:10.1177/03331024221128278
11. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. *J Adv Nurs*. 2000;32(4):1008-1015. doi:10.1046/j.1365-2648.2000.01567.x
12. Barrett D, Heale R. What are Delphi studies? *Evid Based Nurs*. 2020;23(3):68-69. doi:10.1136/ebnurs-2020-103303
13. Ferrari MD, Goadsby PJ, Burstein R, et al. Migraine. *Nat Rev Dis Primers*. 2022;8(1):2. doi:10.1038/s41572-021-00328-4
14. Kerr NL. HARKing: hypothesizing after the results are known. *Pers Soc Psychol Rev*. 1998;2(3):196-217. doi:10.1207/s15327957pspr0203_4
15. Glasser MF, Sotiropoulos SN, Wilson JA, et al. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage*. 2013;80:105-124. doi:10.1016/j.neuroimage.2013.04.127
16. Esteban O, Markiewicz CJ, Blair RW, et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat Methods*. 2019;16(1):111-116. doi:10.1038/s41592-018-0235-4
17. Tanveer S, Alai S, Esterlitz J, Oshinsky M. *Accelerating Headache Clinical Research Studies: The National Institute of Neurological Disorders and Stroke (NINDS) Headache Version 2.0 (V2.0) Common Data Elements (CDEs) (P3.100)*. AAN Enterprises; 2018.
18. Yan CG, Wang XD, Zuo XN, Zang YF. DPABI: data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*. 2016;14(3):339-351. doi:10.1007/s12021-016-9299-4
19. Moayedi M, Salomons TV, Atlas LY. Pain neuroimaging in humans: a primer for beginners and non-imagers. *J Pain*. 2018;19(9):961.e1-961.e21. doi:10.1016/j.jpain.2018.03.011
20. Goldfarb MG, Brown DR. Diversifying participation: the rarity of reporting racial demographics in neuroimaging research. *Neuroimage*. 2022;254:119122. doi:10.1016/j.neuroimage.2022.119122
21. Flanagin A, Frey T, Christiansen SL, Bauchner H. The reporting of race and ethnicity in medical and science journals: comments invited. *JAMA*. 2021;325(11):1049-1052. doi:10.1001/jama.2021.2104
22. Marek S, Tervo-Clemmens B, Calabro FJ, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature*. 2022;603(7902):654-660. doi:10.1038/s41586-022-04492-9
23. Eklund A, Nichols TE, Knutsson H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc Natl Acad Sci USA*. 2016;113(28):7900-7905. doi:10.1073/pnas.1602413113
24. Soares JM, Magalhães R, Moreira PS, et al. A Hitchhiker's guide to functional magnetic resonance imaging. *Front Neurosci*. 2016;10:515. doi:10.3389/fnins.2016.00515