Deliberation 2023-029 of March 23, 2023 National Commission for Computing and Liberties Nature of the deliberation: Other authorization Legal status: In force Date of publication on Légifrance: Wednesday April 05, 2023 Deliberation n°2023-029 of March 23, 2023 authorizing public assistance – Hôpitaux de Paris to implement automated processing of personal data for the purpose of a study on the epidemiology of rare systemic inflammatory diseases in France, entitled “EMIR-Algo” (Application for authorization no. 922278) Commission nationale de l'informatique et des libertés, Seizure on December 2, 2022 by the Assistance Publique – Hôpitaux de Paris of a request for authorization concerning the automated processing of personal data for the purpose of a study relating to epidemiology rare systemic inflammatory diseases in France, entitled "EMIR-Algo"; Having regard to Regulation (EU) 2016/679 of the European Parliament and of the Council of April 27, 2016 on the protection of individuals with regard to the processing of personal data personal nature and the free circulation of such data, and repealing Directive 95/46/EC (general data protection regulations); files and freedoms, in particular its articles 66, 72 and following; Having regard to the favorable opinion of the Ethics and Scientific Committee for research, studies and evaluations in the field of health of June 2, 2022; Having regard to the file and its supplements ;On the proposal of Ms. Valérie PEUGEOT, commissioner, and after having heard the observations of Mr. Damien MILIC, deputy government commissioner,Formulates the following observations:On the data controllerThe data controller for this study is Public Assistance – Paris Hospitals (AP-HP). On subcontractors The Health Data Platform (PDS) will be involved in the implementation of this study. traces, the management of alerts and incidents as well as the management of anonymous data exports, must be formalized by an agreement between the two parties in accordance with Article 28 of the General Data Protection Regulation (GDPR). On the purpose of the processing and its nature of public interest The purpose of the processing envisaged is the implementation of a study entitled "Emir-Algo" relating to the epidemiology of rare systemic inflammatory diseases in France. It should make it possible to develop and validate algorithms for the identification in the National Health Data System (SNDS) of rare systemic autoimmune and auto-inflammatory diseases (BUT), for each major group of diseases or, if this is more relevant, for a specific rare disease. The Commission notes that the "EMIR-Algo" project was one of the winners of the call for expressions of interest launched by the PDS in 2021 relating to the development of algorithms targeting in the SNDS. As such, it will be supported and funded by the PDS. In addition, the targeting algorithms developed or validated as part of this project will be shared with the scientific community through an open source publication for possible reuse. .Subject to the exercise of the right of opposition, three cohorts will be formed as part of this study: a first cohort comprising the data of patients with MAIS included

in the National Rare Disease Data Bank (BNDMR), i.e. approximately 30,000 people, linked with SNDS data; a second control cohort formed from SNDS data, made up of approximately 60,000 non-sick people, aligned in age and sex with the first cohort (two controls for one case); a third cohort comprising the data of patients who do not have a MAIS included in the BNDMR and who have a registration number in the national identification directory of natural persons (NIR) in the BaMaRa application (web application that allows professionals from rare disease reference and skills centers to collect and use their rare disease data themselves), i.e. approximately 48,000 people, linked with SNDS data. These three cohorts will be used to study the performance of established diagnostics by the algorithms for targeting rare MAIS in the SNDS, developed within the framework of this study. secondly, using learning methods on the study data.The purpose of the processing is determined, explicit and legitimate, in accordance with Article 5.1.b of the GDPR, and this processing has a purpose of public interest, in accordance with article 66. I of the law n° 78-17 of January 6, 1978 modified (law "data-processing and freedoms"). On the lawfulness of the processing and the conditions allowing the processing of data concerning health The processing implemented by the AP-HP of Paris is necessary for the execution of the mission of public interest with which it is invested. This processing is, as such, lawful under Article 6.1.e of the GDPR. In addition, this processing, necessary for scientific research purposes, also fulfills the condition provided for in Article 9.2.j of the GDPR allowing the processing of data concerning health. This research project is subject to the provisions of Articles 44.3°, 66. III and 72 et seq. of the "Informatique et Libertés" law, from which it follows that in the absence of compliance with a reference methodology, processing for the purposes of research, study or evaluation in the field of health can only be implemented after authorization from the Commission. On the points of non-compliance with the reference methodology concerned The application file mentions that the treatment envisaged complies with the provisions of the reference methodology MR- 004, with the exception of the nature of the data processed and the procedures for informing the persons concerned. Apart from these exceptions, this processing must comply with the framework provided for by the reference methodology MR-004. On the reuse of existing database data The data from the BNDMR having been authorized by the CNIL (deliberation n° 2019-113 of September 5, 2019; request for authorization n° 2211418) will be reused in the context of this study. These data will be matched with SNDS data as part of the creation of cohorts 1 and 3 of the "EMIR-Algo" study. The Commission notes that the data of the BNDMR are intended to integrate the database of the SNDS catalog pursuant to Article 3 of the decree of 12 May 2022 relating to the data feeding the main database and the databases of the SNDS catalog . On the special categories of data processed With regard to the processing

of SNDS data: Provided that they can be disseminated by the National Health Insurance Fund (CNAM), the data controller requests access to data from the national health insurance system. inter-scheme information from health insurance (SNIIRAM) and the program for the medicalization of information systems (PMSI) for the years 2010 to 2021. Only data that is strictly necessary and relevant to the objectives of the treatment will be transmitted by the CNAM ; in this respect, the filtering and matching of data will be carried out upstream of this transmission by the CNAM. Since the data from the SNIIRAM and the PMSI come from the databases making up the SNDS, all the legislative provisions (articles L. 1461-1 to L. 1461-7 of the Public Health Code – CSP) and regulations relating to the SNDS are applicable in this case, in particular: the prohibition on using this data for the purposes described in Article L. 1461-1 V of the CSP; compliance with the safety baseline applicable to the SNDS provided for by the decree of 22 March 2017. With regard to the procedures for matching BNDMR data with SNDS data: Several procedures for matching data from the BNDMR with SNDS data are provided in the application file: initially, a deterministic matching will be carried out using the NIR, gender and full date of birth of patients appearing in the BaMaRa application. The NIRs will be sent to the CNAM directly from the BaMaRa database, by the team in charge of this database, which is separate from that in charge of the BNDMR. for patients whose data could not be matched deterministically , a probabilistic matching will be carried out using the date and place of treatment of the patient in the expert network, their date of birth (month and year), their place of residence (municipality and department), their sex and processing or diagnostic codes. With regard to the matching method based on the NIR, the Commission draws the attention of the data controller to: the necessary effectiveness of the partitioning between the teams working on the BNDMR and on BaMaRa, of which he must ensure by organizational and technical measures adapted to the risks and by regular checks; the need for the "Operations" division, which operates BaMaRa and will carry out the extraction of the NIRs, to delete the correspondence table [IDMR – IDacc] which is transmitted to it as a parameter of the extraction program; this deletion must take place at the end of the extraction procedure, the table being kept only by the BNDMR team; the need for each extraction request to use a different series of attachment numbers (IDacc), to limit the risks of cross-re-identification. The Commission considers that the data whose processing is envisaged are adequate, relevant and limited to what is necessary with regard to the purposes of the processing, in accordance with the provisions of Article 5.1.c of the GDPR. On information and the rights of individuals With regard to participants who received an individual information note when the BNDMR was set up, including a specific information system to which they can refer prior to the implementation of each new study carried out using data from the

warehouse: An information note relating to this study will be distributed before the start of the study on the website dedicated to the BNDMR. It must include all the information provided for by the GDPR. With regard to the other participants in the study: Pursuant to article 69 of the law "Informatique et Libertés" and article 14.5.b of the GDPR, the obligation to provide individual information to the data subject may be subject to exceptions, in the event that the provision of such information proves impossible, requires disproportionate effort or seriously compromises the achievement of the objectives of the processing. In such cases, the controller shall take appropriate measures to protect the rights and freedoms, as well as the legitimate interests of the data subject, including by making the information publicly available. In this case, an exception will be made to the principle of individual information of persons and appropriate measures will be implemented by the dissemination of collective information on the website of the data controller. This information will also be disseminated by display in the rare disease reference and competence centers as well as on their website. This processing will also be recorded in the PDS transparency portal. These information methods are satisfactory to the with regard to the provisions of the GDPR and the law "Informatique et Libertés". With regard to the procedures for exercising rights:The persons concerned may exercise their rights with the data protection officer of the data controller throughout the duration of the study.These procedures for exercising rights are satisfactory with regard to of the provisions of the GDPR and the "IT and freedoms" law. On accessors and recipientsOnly the data controller and the persons authorized by him have access to the data within the framework of this authorization. The data controller keeps documents up to date indicating the competent person(s) within it to issue the authorization to access the data, the list of authorized persons, their respective access profiles and the methods of allocation, management and control of authorizations. These categories of persons are subject to professional secrecy under the conditions defined by Articles 226-13 and 226-14 of the Criminal Code. The qualification of authorized persons and their access rights must be regularly reassessed, in accordance with the methods described in the authorization procedure established by the data controller. On the security of the data and their hosting methods As a preliminary point, the Commission notes that the application file justifies the need to use the technical solution of the PDS, taking into account the characteristics as well as the specific methods of implementation of this study, in particular the use of a set of software and technical solutions made available by the PDS. The security of the data of the project space dedicated to the "EMIR-Algo" project depends essentially on the technical solution of the PDS, which has been the subject of an overall analysis of the risks and the impact on privacy, followed by an approval according to the SNDS security reference system .More specifically, an impact analysis relating to data protection has been

sent to the CNIL concerning the technical solution of the PDS, which corresponds to a secure SNDS bubble and which will host the "EMIR-Algo" project. The data controller has carried out and transmitted in support of the authorization request an impact analysis relating to data protection specific to the "EMIR-Algo" project and integrating the elements provided by the PDS for its technical solution. An approval of the project space was thus carried out by the data controller on October 24, 2022, for a period of three years, subject to the implementation of the action plan that it defined. approval is only valid until October 24, 2025 and must therefore be renewed before this date if the project is still in progress. The security measures implemented by the data controller appear proportionate to the risks presented by the processing. transfers of data outside the European Union The provisions of article R. 1461-1 of the CSP provide that no transfer of personal data can be carried out outside the European Union, except in the case of occasional access to data by persons located outside the European Union, for a purpose falling under 1° of I of Article L. 1461-3 of the CSP. In this case, the application file mentions that, although the service provider is not exclusively subject to the laws and jurisdictions of the European Union, no transfer outside the European Union of individual data from the SNDS is planned, no member of the research team being located outside the European Union. On the retention period of the data The data necessary to carry out the deterministic matching (in particular the NIR) will not be kept after the matching The other study data will be made available for three years. These data retention periods do not exceed the periods necessary for the purposes for which they are collected and processed, in accordance with the provisions of Article 5.1.e of the GDPR.Additional observationsAs regards models developed by automatic learning methods, a risk study must be carried out in order to verify that it is not possible to extract personal data from them or to deduce the presence of a person in the training dataset, prior to any publication in open source. Authorizes, in accordance with this deliberation, the Assistance Publique – Hôpitaux de Paris to implement the aforementioned processing. President Marie -Laure DENIS