Response to the NHA Consultation Paper on the Proposed Health Data Retention Policy

Submission from IT for Change

December, 2021

IT for Change

Consultation on the
National Health Authority's Proposed Health Data Retention Policy
Response submitted by IT for Change[1]
December, 2021

Summary of our recommendations:

1. **Restrict the ambit of the retention policy, for the time being, only to ABDM registered facilities.**
2. **The ultimate aim should be to cover all health facilities, so as to ensure appropriate representivity in data of all sections of the population, and avoiding data biases.**
3. **Small health facilities need to be provided support for employing certified services for health data storage.**
4. **There must be absolute clarity about which actors are covered or intended to be covered by the policy, and which not.**
5. **Individuals must be able to 'opt-out' of sharing of data, even in anonymized forms, by health facilities with third (or outside) parties, unless involving legal requirements.**
6. **Differentiated retention periods may be mandated on the basis of a more granular classification, not just as per different kinds of data, but also the kinds of users of health data, and the nature of potential data use.**
7. **Anonymization of the retained data should not be allowed to be employed for indiscriminate and/or otherwise inappropriate data sharing and expropriation, particularly as involving commercial actors.**

Released by the National Health Authority (NHA) for public consultation on 23rd November 2021, the Proposed Health Data Retention Policy aims to define a minimum period of retention of data held in 'Health Facilities' registered under the Ayushman Bharat Digital Mission (ABDM). These include both public and private health facilities including hospitals, clinics, diagnostic laboratories and imaging centers, pharmacies, etc[2].

1 www.ITforChange.net
2 Health Facility Registry

IT for Change appreciates the proactive seeking of feedback in building out the framework for governance of health data under the ABDM. We agree that the retention of medical records is a critical practice with many potential long term benefits, as highlighted by the experiences of developed nations that have adopted these policies and rules.

BACKGROUND AND RELEVANT ISSUES

Objectives of health data retention: Retention policies for medical records are designed to serve two broad functions:

1 For individuals, the retained records help create an accessible and interoperable longitudinal view of a patient's medical history. These records help health professionals better service patient needs, and are also useful to defend a complaint or claim of clinical negligence, forensic cases, and for processing insurance claims.

2 For society more broadly, the retention of records help facilitate research on and new developments for improved healthcare delivery – through cross-sectional studies within a given hospital or longitudinal studies on geographically distributed patients, among others[3]. The shift from industrial age health systems and services to digital age ones is anchored on data-based services, care, medicines, etc. Even if it has to go through a considerable transition period, we cannot neglect in ensuring that the new digital health systems, as and when they mature, are inclusive and beneficial to all.

A third layer, between the individual and society, is that of a specific group, community etc. that an individual whose data is retained may belong to. Such a group/ community may have common and unique health-relevant features, also expressed in their data. There could accordingly be group/community specific health related benefits and harms as well, an area which is yet less recognized, but it is beginning to be[4]. Any health data retention policy must also take note of this.

Downsides of health data retention: On the downside, more collection and longer retention of health data related to individuals and groups creates new challenges of data security and data privacy. A health data retention policy must in any case be applied only after a data protection law is in place, which is expected soon. A further issue, as mentioned, is about possible collective harm, including commercial misappropriation.

Creation of new data-sets, and corresponding challenges: Data retention policies may result in creation of new digital data records where they may not have existed before. Such new facilities

3 Using Electronic Health Records for Population Health Research: A Review of Methods and Applications | Annual Review of Public Health

4 See in this regard the draft report of the Committee of Experts on Non-Personal Data Governance Framework https://static.mygov.in/rest/s3fs-public/mygov_159453381955063671.pdf
Also some Covid pandemic related documents like WHO linked initiative 'Access to Covid-19 Tools Accelerator's' Personal Data Governance Framework https://www.finddx.org/wp-content/uploads/2021/01/ACT-A-Dx-data-governance-framework_15.01.2021.pdf

undertaking electronic data retention may not have adequate capacity and skills, or simply inclination, for data protection. This calls for new and heightened data security and data privacy related vigil, regulation, and enforcement. Such regulation and enforcement currently exists only for personal data (assuming a data protection law in place), which are important to apply ever more diligently. Meanwhile, greater data retention obligations creating more electronic health data records also further accentuates the issues of security of anonymized, non-personal data, and collective privacy. It also makes it more important and urgent to develop benefit-sharing frameworks for such anonymized aggregate data vis-a-vis the collective data subjects, which could be a group, community, a nation or the whole global public.[5]

New grounds for non-consensual data storage: Health data retention policy potentially creates grounds for a data collector/holder to cite legitimate purpose, public interest purpose or reasonable purpose, to retain data even without consent. Even under EU's General Data Protection Regulation, consent is not necessary for many kinds of processing of data, like for "the purposes of preventive or occupational medicine, medical diagnosis, provision of health or social care or treatment, management of health or social care systems and services, under a contract with a health professional or another person subject to professional secrecy under law (the 'medical care' ground)".[6] Worldwide, legally enforced health data retention rules/guidelines have meant that the individual concerned may not even be able to seek deletion of her data, or prevent many other kinds of 'legitimate processing' of her data. The statements and documents of the government are quite ambiguous on this issue, but this is a very likely/possible scenario. It is for this reason that it should be absolutely clear which 'actors' come under the proposed health data retention policy, and which not, thereby clarifying who can take or claim the 'non-consensual retention' cover that the policy may provide and who cannot. In this respect, an open-ended definition and listing of 'health facilities', with the use of 'etc.' in the end, is not at all appropriate.

Shifting notions of health services, and the question of policy's coverage: An important emerging issue is the proliferation of digital platforms and apps that in many ways undertake activities and provide services similar to those of traditional 'health facilities'. For instance, Practo-like platforms collect a lot of transactional data (which kind of data is all-important in today's platform economy), similar to what a 'health facility' would normally collect about medical appointments, consultations, etc. Wellness devices, like Fitbit watches, are not very different from wearable diagnostic tools used by 'health facilities' for instance for ambulatory blood pressure monitoring. This may, on the one hand, lead to health data regime shopping (being able to consider oneself subject to different regimes – or not – as per convenience and benefit) and, on the other hand, result in application of different regimes

5 See in this regard the draft report of the Committee of Experts on Non-Personal Data Governance Framework https://static.mygov.in/rest/s3fs-public/mygov_159453381955063671.pdf

6 The Final GDPR Text and What It Will Mean for Health Data - Chronicle of Data Protection

for similar activities and actors, depending on whether they are undertaken by covered 'health facilities' or others. Both these scenarios need to be avoided.

Low levels of digitalization, and low capacity: An India-, or developing countries-, specific issue is the low level of digitization and low public sector spending on health. India is on the lower end of the spectrum of countries spending on health care. Its public health spending is only about 1.3% of its GDP[7]. It also has fewer hospital beds and physicians for its population – only 8.5 beds and 8 physicians per 10,000 people – compared to countries like South Korea and Japan.[8] Levels of digitalization of health facilities is also very low, particularly outside big urban centres. Excessive and indiscriminate pushing of digital data retention mandates in such conditions may lead to unacceptable levels of hazards in relation to data security, data privacy, and data expropriation.

Need for adequate data representation: On the other hand, it is important to aim towards universalising digital health data maintenance and retention across the country, reaching all its nooks and corners, so that all sections of the population (1) have the benefit of longitudinal data views for better health care and outcomes, and (2) get represented in health data sets at regional, national, and global levels as more and more developments in health care becomes data and AI based.

OUR RECOMMENDATIONS

Any health data retention policy needs to appropriately balance across all the factors and considerations discussed above. We offer the following recommendations for the proposed health data retention policy:

1. **Restrict the ambit of the retention policy, for the time being, only to ABDM registered facilities.** Smaller 'health facilities', especially in distant rural and semi-urban areas, may lack the technical capacity required to comply with a digital health data retention policy. Even if they overextend themselves to comply, it may open up considerable hazards in relation to data security and privacy. It is therefore prudent to take a phased or graduated approach for extending the coverage of a digital data retention policy. This can be done by restricting the initial application of the policy only to the ABDM registered facilities.
2. **The ultimate aim should be to cover all health facilities, so as to ensure appropriate representivity in data of all sections of the population, and avoiding data biases.** Given the long term significance of digital health data retention, it is important that the government starts to take strong immediate steps towards encouraging and nudging the broader ecosystem towards its adoption. Applying the policy first to ABDM registered 'health facilities' can in various ways become an incentive for 'health facilities' across the country to undertake appropriate digitalization. This should be further pushed through awareness and capacity

7 Healthcare sector: Hits & misses from budget 2021 - Times of India
8 India's economy needs big dose of health spending - Livemint

building programs in collaboration with institutions such as the Indian Council of Medical Research and the National Medical Commission.

3 **Small health facilities need to be provided support for employing certified services for health data storage.** This should be ensured in a manner that is secure and enables full ownership and control over its data by the 'health facilities' concerned. The National Health Authority (NHA) may release model documents to guide the procurement of such storage services, while supplementing the budget of public health facilities for an initial period to account for the expense. The data storage services should be able to exercise no independent access to or use of data, even of anonymized data, or metadata, other than as strictly required for providing the storage and access service. The NHA may, in addition, organize workshops to build the skills necessary for small facilities to negotiate terms with data storage and access providers. It may also create an empanelled list of such providers that have been audited for data security.

4 **There must be absolute clarity about which actors are covered or intended to be covered by the policy, and which not.** Health data retention policies will need to remain aware of the emergent situation whereby, as discussed, the line between traditional 'health facilities' and services, and new kinds of digitally-enabled ones is blurring.So, we recommend that data retention policies should at present only be applicable to clearly defined 'traditional' 'health facilities', with no scope of definitional ambiguity. This is because a data retention policy can provide, and thus be cited as, grounds for non-consensual storage and processing of data. It is possible that, at a later stage, more activities and actors are found necessary to be brought under such a policy in order to fulfil the basic objectives of data retention policies mentioned at the beginning of this paper. Any such inclusion should follow a due diligence process, giving the matter an all-round consideration, including through appropriate consultations. If found fit to be covered under the data retention policy, such activities and actors must simultaneously be made subject to all other laws, rules, guidelines and standards that apply to 'traditional' health facilities, given the sensitive nature of health services.

5 **Individuals must be able to 'opt-out' of sharing of data, even in anonymized forms, by health facilities with third (or outside) parties, unless involving legal requirements.** Even as the data retention policy/guidelines are applied, special arrangements must be created for data pertaining to individuals who are unwilling to share their data horizontally for any purposes, including research, whether in raw, anonymized or pseudonymized form. For example, UK's National Health Service (NHS) launched a national 'opt-out' program in 2018 that created a single opt-out point applied across the system for patients unwilling to have their data shared outside the NHS for purposes of research planning, and provided a

mechanism for people to register their choice[9]. Opt-out provision may specifically or i*nter alia* include refusing anonymization and aggregation of one's health data. There could also be cases where patients do not wish for their data to be used at all, even within a system, but the same cannot be deleted as per the retention policy. In such cases, options may be considered for enabling such data to be held in a digital 'lockbox'. For instance, in Ontario, Canada, the 'lockbox provision' enables a health facility to put patient data in a sealed envelope for the duration of retention when the patient wants the data deleted, which thereupon cannot be disclosed without consent, or if otherwise required by law.[10] In the Indian context, this is visualized within the Health Data Management Policy whereby stored personal data can be blocked or restricted in cases where the law prohibits deletion.[11]

6 **Differentiated retention periods may be mandated on the basis of a more granular classification, not just as per different kinds of data, but also the kinds of users of health data, and the nature of potential data use.** Granularity of data ensures that the key attributes, outputs, and trends can be monitored effectively[12]. The retention periods for different granular classifications may be modified to reflect the priorities and objectives of the retention policy. This ensures that data that is less important to retain is not over-retained, and that which is more important is not under-retained. For example, data that may be used for research purposes (such as cancer diagnostic images) should have a higher retention period with adequate safeguards against misuse. As more and more kinds of data become relevant to health, the understanding of health data is expanding, as with the blurring lines between traditional 'health facilities' and new kinds of health services. These include personal wellness data (such as feeds available from wearables, smart watches, fitness machines and numerous diet, exercise and social apps), proxy data (ranging from Facebook likes and Instagram comments to location and environmental data, resident postcodes and even bathroom and fridge access) and self reported data (such as blood pressure, heart rate, glucose levels, temperature, weight and in-home remote monitoring)[13]. The proposed health data retention policy should therefore be very clear what kinds of data are included in its remit, and what are not. If more kinds of data are in future to be included under the mandate of health data retention policy, a specific process has to be set up for it, with all due diligence and giving consideration to all the relevant facts. Any such new inclusions should be fully justified, and involve an adequately consultative process. While ensuring individual and social benefits of

9 National data opt-out - NHS Digital
10 Retention of Personal Health Records - College of Respiratory Therapists of Ontario
11 Clause 14.1(b)(ii)(D), Health Data Management Policy
12 Reimagining Data Architecture | NITI Aayog
13 Future of Patient Data | Accenture

data-based healthcare is important, the importance of data minimization principle needs to equally be kept in mind.

7 **Anonymization of the retained data should not be allowed to be employed for indiscriminate and/or otherwise inappropriate data sharing and expropriation, particularly as involving commercial actors.** New data retention obligations give rise to newer and deeper concerns about how the retained data gets used, abused and expropriated. Applying a data retention policy would lead to greater retention of health data, and among new actors that may not earlier have maintained digital health records. Whereas this brings up issues of data security and privacy, there is also the, under-examined, issue of how anonymized, non personal data, obtained from these records gets used, abused and expropriated. This is especially so because anonymized, non-personal data still does not come under any legal coverage (even if we assume a data protection law in place). Various data subject groups can get exposed to collective harms – for instance, as involving a community of patients with a rare disease, or an ethnic group having some unique genetic or other health reared data. There is also the issue of what commercial and other benefits are derived from such anonymized, non personal data, and the right of the collective data subjects concerned to a share in such benefits. The second draft report of the Committee of Experts on Non Personal Data Governance Framework[14] recommends legal provisions both for prevention of collective harm (through putting a legal obligation of 'duty of care' on collectors of non personal data, and enabling a right of any member of the community/group concerned to take to the court any complaint of collective harm). It also provides legal basis as well as means for benefit sharing in relation to use of non personal data related to a group or community. These concepts and proposed provisions are ever more relevant for health related anonymized, non-personal data, and should be applied to it. Any data retention policy should be accompanied by appropriate and adequate laws, rules and policies aimed to prevent individual as well as collective harm from processing not just personal data but also anonymized non personal data. They must also enable mechanisms for sharing the benefits arising from any such processing. Since the health sector is tightly regulated, many such measures for protections against data-related harm – individual as well as collective, and sharing of benefits from data use, can be ensured through health sector specific rules, codes and guidelines.

14 Report by the Committee of Experts on Non-Personal Data Governance Framework