Society of Medical Education & Research

2023, Volume 3, Page No: 115-125
ISSN: 3108-5059

Asian Journal of Ethics in Health and Medicine

Exploring Challenges and Enablers in Applying Automated Decision Support Systems for Genomic Data Access: A Qualitative Interview Study

Stephanie Best[1,2*], Jerrin S. Long[1], Faris Leadbeatter[3,4], Amit Braithwaite[1]

[1] Australian Institute of Health Innovation, Macquarie University, Sydney, NSW 2113, Australia.
[2] Australian Genomics Health Alliance, Murdoch Childrens Research Institute, Melbourne, VIC 3052, Australia.
[3] Melbourne Genomics Health Alliance, Walter and Eliza Hall Institute, Melbourne, VIC 3052, Australia.
[4]Department of Paediatrics, University of Melbourne, Melbourne, VIC 3010, Australia.

***E-mail** ✉ Jarrinslong@outlook.com

Abstract

Data access committees (DAC) serve as gatekeepers for secured genomic and associated health datasets but face increasing difficulty in managing the growing scale and complexity of data production. Automated decision support (ADS) tools have demonstrated potential in enhancing consistency, compliance, and coordination during data access review processes. Despite this, little is known about how DAC members themselves view the added value of ADS in improving the quality and efficiency of their decision-making. This qualitative research draws on 13 semi-structured interviews with DAC members across multiple regions to explore barriers and facilitators to the adoption of ADS in genomic data access management. Participants expressed general support for pilot testing of ADS functions, such as classifying data types, validating user identities, and labeling datasets with usage terms. Nonetheless, reservations emerged around risks of excessive automation, insufficient human oversight, limited institutional prioritization, and potential conflicts with organizational missions, which tempered enthusiasm. Institutional pressures for change, alongside perceived relative advantages of ADS compared to current practices, strongly influenced DAC members' considerations of implementation. Further investigation is required to strengthen the evidence base regarding the comparative effectiveness and decision outcomes between institutions that integrate ADS into their workflows and those that do not.

Keywords: Ethics, Data access committee, Automation, Implementation science, Genomic data, Decision support

Introduction

Genomics has become one of the most data-intensive scientific domains, with projections indicating that by 2025 it will generate more storage and computational demands than platforms such as Twitter, YouTube, and even the field of astronomy combined [1]. To address the accelerating need for efficient collection, management, and use of genomic data, large-scale national genomics initiatives [2] have increasingly turned to centralized repositories that encourage data pooling and incentivize data sharing [3–5]. Many of these initiatives have adopted the data commons model [6], which prioritizes collaborative research and open data access over proprietary restrictions [3]. Within this ecosystem, data access committees (DACs) hold primary responsibility for ensuring that only bona fide researchers, pursuing projects permitted by participants' informed consent, are approved for access [7]. DAC membership often consists of compliance officers, researchers, and at times, data security specialists, serving either in paid roles or as volunteers. At its most basic, the function of a DAC is to adjudicate requests, granting access provided minimum standards of compliance and data protection are satisfied.

Access this article online https://smerpub.com/

Received: 21 August 2023; **Accepted:** 18 November 2023

How to cite this article: Best S, Long JS, Leadbeatter F, Braithwaite A. Exploring Challenges and Enablers in Applying Automated Decision Support Systems for Genomic Data Access: A Qualitative Interview Study. Asian J Ethics Health Med. 2023;3:115-25. https://doi.org/10.51847/QM2eCJ69fc

Growing recognition of privacy risks and criticisms of DACs' focus on compliance alone, however, have prompted debate over whether their responsibilities should also extend to evaluating the broader social and scientific value of proposed projects [8]. Our previous empirical research [9] revealed ongoing disagreement regarding this expanded scope of DAC oversight, particularly where it overlaps with ethical considerations traditionally overseen by institutional review boards.

For instance, Cheah and Piasecki argue that DACs must balance the imperative of facilitating data sharing with safeguarding the rights of individuals and communities represented in the datasets. They suggest that "data access should be granted as long as the data reuse fulfils the criterion of having even a minimal social value, and minimal risk to data subjects and their communities" [7]. In this way, DACs play a pivotal role in maintaining responsible data sharing environments, as they govern access to genomic and, increasingly, other forms of health data [10–12].

Despite this, DACs are not always regarded as the most efficient means of managing data access compared to other review mechanisms [13]. Under the prevailing model, DACs manually evaluate each application, judging it against criteria such as whether the requested data are appropriate, whether the intended use complies with provider stipulations, and whether institutional and legal requirements for data privacy and security are met [7]. As with many human-driven processes, this review can be slow, labor-intensive, and susceptible to errors. For example, DAC members may interpret data use terms differently, particularly when such terms are ambiguously worded [14]. These interpretive discrepancies may lead to subjective judgments about whether a request genuinely fits within the permitted scope of use. Furthermore, variations in how consent form conditions are written, interpreted, and applied across different DACs contribute to inconsistencies and delays in data access, raising the risk of misalignment with the original terms under which participants contributed their data [14].

Beyond the review stage, other aspects of the data access pipeline can also generate bottlenecks. Evidence suggests that the manual review and execution of data access agreements is increasingly inefficient, inconsistent, and prone to mistakes [13,15]. In addition, many researchers still depend on the outdated copy-and-download method for accessing data once approval is granted. This approach not only heightens security vulnerabilities [11] but is becoming untenable given the sheer size and complexity of genomic datasets [16,17].

To address these challenges, standards developers and software engineers have worked to introduce semi-automation into three core components of cloud-based data access control: user authentication, review of access requests, and verification of alignment between proposed research and data use terms [14]. Automated decision support (ADS) systems represent one such strategy, integrating algorithms, ontologies, and software [18] to support the classification, storage, and operationalization of decisions in the access review process. The Data Use Oversight System (DUOS) exemplifies such an ADS tool [19]; in recent pilot evaluations, DUOS achieved full agreement with human DAC decisions [15] and was able to codify 93% of genomic datasets in NIH's dbGaP [20]. Although ADS solutions can augment DAC operations with semi-automated review processes, systematic evidence is still lacking regarding the barriers and facilitators that influence their implementation in practice [21]. In particular, we know little about how DAC members themselves assess the value added by ADS—whether in terms of efficiency, quality, or decision accuracy—and what organizational challenges they foresee in adopting these systems given the diversity of institutional arrangements under which DACs function.

The convergence of ADS development with large-scale cloud migration, which promises near-instantaneous access approvals, makes this a timely moment to investigate implementation challenges and opportunities. The genomics community can also draw lessons from earlier attempts to deploy ADS technologies in fields such as public health [22], law enforcement [23], and clinical medicine [24], where insufficient stakeholder engagement hindered success. In this study, we present empirical findings on the "constellation of processes" that shape the adoption of ADS for genomic data access management and propose actionable recommendations for institutional data stewards considering, or already undertaking, such implementation.

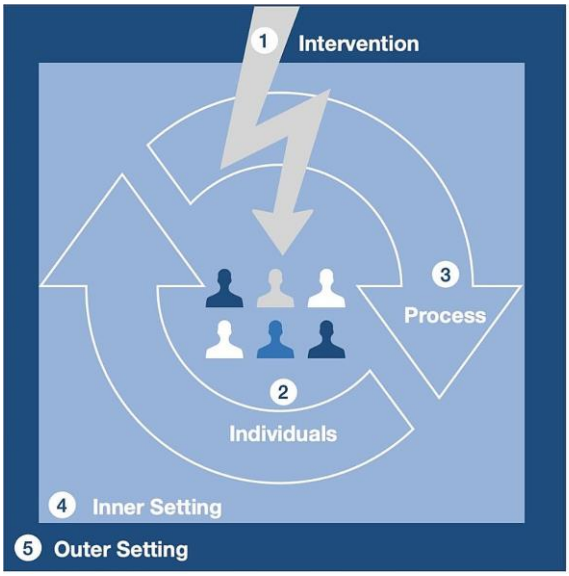
Methods

This study employed a qualitative descriptive approach to engage prospective end users of automated decision support (ADS) in genomic data governance. Our aim was to examine two central questions: What barriers and facilitators shape the implementation of automated workflows for managing genomic data access requests,

and how might ADS influence the quality and effectiveness of data access committee (DAC) reviews? To frame our investigation, we adopted Damschroder and colleagues' conceptualization of implementation as the "critical gateway between an organizational decision to adopt an intervention and the routine use of that intervention" [25]. This perspective emphasizes studying the "constellation of processes intended to get an intervention into use within an organization" [25].

We applied the Consolidated Framework for Implementation Research (CFIR) to compare procedures and workflows governing genomic data access, with the goal of better understanding how automated systems could be integrated across international, publicly funded genomic repositories. The CFIR outlines a comprehensive set of constructs organized into five key domains of effective implementation, developed through rigorous meta-theoretical synthesis of diverse implementation models **(Figure 1)**. Beyond its theoretical grounding, CFIR serves as a practical tool for systematically identifying potential facilitators and obstacles prior to adopting a new innovation [25]. Its adaptability also makes it well suited for surfacing bioethical dimensions of implementation in genomics, as demonstrated in prior studies.

An interview guide, tailored specifically for this research, was developed and is provided in Supplementary Materials 2.



CFIR Domain	Application to DUOS implementation [relevant CFIR constructs]
Intervention Characteristics	**Is DUOS superior to the status quo? Can it be adapted to work across DACs?** [Intervention source, evidence strength, relative advantages, adaptability, trialability, complexity, design quality and cost]
Individuals	Do DAC members **posses both the capability and motivation to use DUOS?** [Knowledge, attitudes and beliefs, self efficacy]
Process	**What existing workflows does DUOS affect? Whose buy-in, input and expertise is needed?** [Planning, engaging, executing, reflecting, evaluating, external change agents, champions, opinion leaders]
Inner setting	**What is the feasibility of embedding DUOS into the DAC review system?** [Structural characteristics, networks and communication, culture, implementation climate (e.g. organizational readiness, leadership, available resources, intervention knowledge and access)]
Outer setting	**What gap does DUOS address and why should we implement now?** [Needs and resources, cosmopolitanism, peer pressure, external policy & incentives]

Figure 1. The Consolidated Framework for Implementation Research (CFIR), adapted for this study, outlines five core domains—Intervention Characteristics, Individuals, Process, Inner Setting, and Outer Setting—that informed how 13 qualitative interviews were organized to examine influences on the adoption of automated decision support (ADS) in genomic data access oversight and sharing across publicly funded repositories worldwide.

Data collection

From 27 April to 24 August 2022, we carried out 13 semi-structured interviews that together included 17 data access committee (DAC) members. Many of the participants were drawn from a prior survey study [9], where they had agreed to a follow-up conversation. Additional participants were identified either through the Data Access Committee Review Standards Working Group (DACReS WG), chaired by VR, JL, and ESD, or through online searches targeting publicly funded genomic data repositories. All interviews were hosted virtually via Zoom and both video and audio were recorded. Discussions followed a guide adapted from validated tools in the official CFIR repository (https://cfirguide.org/evaluation-design/qualitative-data/). The guide, tailored to the ADS context, contained 29 questions probing barriers and facilitators in DAC reviews, including dimensions such as Inner Setting, Outer Setting, and Intervention Characteristics. Sessions

lasted approximately 45–60 minutes. The finalized guide is available in Supplementary Materials 2.

Data analysis

Interview transcripts were examined using a framework analysis strategy (Pope, Ziebland, and Mays 2000). A deductive coding template, based on the publicly available CFIR codebook (see Supplemental Materials 1), was applied. To ensure coding reliability, two reviewers (VR and JB) independently tested the schema on three transcripts and achieved an interrater reliability score of 0.83 before extending the process to the remaining interviews. Any discrepancies identified during the pilot stage were resolved through collaborative discussion until agreement was reached.

Results

(a) Geographical, institutional, and demographic background of participants

Of those interviewed, 41% were affiliated with U.S. DACs, while the other 59% represented committees in Canada, the United Kingdom, Spain, Tunisia, Australia, and Japan **(Table 1)**. Regarding institutional settings, nearly 60% of participants worked in non-profit research institutes, 24% were based in academic-affiliated research organizations, 12% were part of government research agencies, and 6% came from research consortia. Gender demographics showed that 76% of interviewees identified as female and 24 percent as male.

Table 1. CFIR Code Application Results

High frequency>=25	Medium frequency10–25	Low frequency< 10
Code (number of code applications)		
Tension for Change (98), Relative Advantage (72), Knowledge &Beliefs about the Innovation (47), Structural Characteristics (36), Planning (33), Cosmopolitanism (32), External Policy & Incentives (30), DAC tools (*code created by the team)* (30), Compatibility (30), Needs & Resources of those Served by the Organization (27), Key Stakeholders (25)	Culture (23), Networks & Communications (22), Relative Priority (21), Cost (21), Adaptability (21), Innovation Source (20), Reflecting & Evaluating (19), External Change Agents (18), Formally Appointed Internal Implementation Leaders (17), Available Resources (16), Individual Identification with Organization (15), Access to Knowledge & Information (14), Evidence Strength & Quality (14), Peer Pressure (12), Other Personal Attributes (12), Opinion Leaders (11)	Individual Stage of Change (9), Goals & Feedback (9), Champions (9), Implementation Climate (8), Engaging (8), Self-efficacy (7), Leadership Engagement (7), Complexity (6), Trialability (6), Learning Climate (6), Characteristics of Individuals (3), Innovation Participants (3), Readiness for Implementation (2), Executing (2), Design Quality & Packaging (1), Process (1)

(a) Opportunities for ADS

Table 2 summarizes how often CFIR implementation factors appeared across the interviews. From this analysis, three primary facilitators emerged in relation to adopting ADS for genomic data governance: (1) external policy pressures combined with the demand for more efficient workflows, (2) institutional capacity to expand and adapt ADS, and (3) the role of interoperability.

Table 2. Participant demographics

DAC location	Total % (n)
United States	54 (7)
Canada	12 (2)
United Kingdom	6 (1)
Spain	6 (1)
Tunisia	6 (1)
Australia	18 (3)
Japan	12 (2)
Institutional type	
Non-profit Research Institute	59 (10)

External policy and need for efficient workflows

Participants described ADS adoption as partly motivated by compliance with emerging data sharing requirements introduced by major research funders such as the National Institutes of Health and reinforced by peer-reviewed journals. These new expectations have expanded the workload of ethics oversight bodies [26], including DACs, which continue to manage their reviews primarily through manual processes [9]. In most cases, committee members examine each request individually, often in the order received, and then issue a decision. With the number of data access requests projected to rise substantially [27], interviewees anticipated that implementing ADS could reduce both costs and workload while maintaining review quality and minimizing risks of noncompliance with data use conditions.

Several participants emphasized that ADS could streamline the intake stage by ensuring submitted requests were consistent with the consent terms originally agreed to at data collection. Many noted that the initial triage of Data Access Requests (DARs)—checking that forms were complete and aligned with the rules for access—was frequently the bottleneck in the submission-to-decision pathway. This stage involves verifying that required details such as study aims, requested datasets, and ethics approval are included, yet requirements vary by institution and dataset, making the process lengthy. To illustrate this variability, we asked participants to share copies of their DAR forms, which ranged from 3 to 18 pages and differed considerably in content. Interviewees agreed that ADS could assist here by automatically identifying missing elements, validating the requester's identity and supporting documents, and notifying applicants if additional information was required. One participant summarized the potential advantage of this approach by stating:

Because one of the biggest concerns in our DAC is that sometimes it takes too much time to be read by all the nine members. ... They're institutional directors or university professors. So I think it will help. Maybe if you have 50% of the work done by an automated system, so you just have to do the 50%. I think ... this will be a good motivation for them saying 'OK' [to implement ADS].

- Participant M.

Scalability and cost effectiveness

Interviewees suggested that ADS-driven workflows offer a scalable and economical approach for handling both newly produced datasets and legacy data once grant support has expired. They noted that ADS systems are capable of efficiently archiving and rapidly retrieving data use conditions while also maintaining records of previous DAC decisions for auditing purposes. In particular, two participants highlighted the persistent difficulties of identifying cost-effective strategies for managing older, legacy datasets:

Actually there are lots of costs related to data sharing, particularly if I'm sharing data from the 1990s, for example. I don't have any money or budget any- more to prepare the data [for secondary uses]. ... And similarly, when it comes to these reports [on data sharing activities], there's no extra money for doing the work to create those reports. But we're having to report back over assets from years, decades in fact. And there was always just a little bit of a hint 'oh well, maybe we'll find some money'. No, no, you have to find it out on your own.

Participant F.

I mean potentially as we grow over the years, you know what's going to happen. ... we've also discussed some scenarios, where, for example, we find ourselves with a larger amount of requests coming in, [and] we only accept applications up to certain days and then, we open this next quarter, close it again. But there potentially could be room for automation depending on the increase in request in the coming years.

Participant A.

Retention and sustainability of human resources

Interviewees emphasized that challenges in maintaining repository staff and DAC membership represent a significant driver for ADS adoption. Several participants pointed out that automated systems could help preserve consistency in review processes when personnel turnover occurs, such as when DAC members or data generators leave an institution. Smaller institutions, unlike large, well-resourced government repositories, often lack sufficient human resources to support long-term oversight, preservation, and management of increasingly complex and voluminous datasets, making ADS an attractive solution to mitigate these gaps:

As the program scales, the participant diversity scales, the data diversity scales. I think it is almost impossible

to see a scenario where we do not rely on some level of automation to support human decision making about what is responsible use.
Participant J.

Interoperability

DAC members indicated that ADS systems have the potential to offer centralized, interoperable frameworks that enhance both inter-organizational and international genomic data sharing. Participants suggested that ADS could encourage the consistent use of standardized request forms, access agreements, dataset identifiers, and procedures for validating researcher identities. One interviewee illustrated this point by stating:
But this [ADS] will free up a lot of time in the process is it also potentially means that it will become easier for, if you're working in a team to hand off tasks as well because you will have a single system. … Also, consistency between organizations. If we have multiple organizations take this up, it's going to mean less lead time. [Let's] say people take a new job in a new place. We'll actually have some software that people will recognize and be able to use and uptake, which we've been trying to go towards without ethics approval processes within the hospital and health services… [standardized] systems makes it easier for actual communication between organizations on processes, because everyone kind of begins to know what's happening.

- *Participant E.*

(b) Barriers to implementing ADS

Although ADS offers clear benefits for genomic data access management, interviewees highlighted several obstacles to its integration within DAC workflows. These included: (1) lower prioritization relative to more urgent governance concerns, (2) insufficiently equipped personnel and institutional infrastructure, (3) financial constraints, and (4) the need for adequate human oversight.

Prioritization

Many participants noted that institutional leaders often focus on other pressing research data priorities rather than investing in new data governance solutions, such as enhancing data quality, expanding diversity in datasets, fostering collaborations with underrepresented researcher and participant groups, and releasing datasets. While researchers generally recognize the importance of thorough and responsible data access review for genomic data sharing, immediate concerns around data quality frequently take precedence. Participants also suggested that ADS adoption may be deprioritized because no major data breach or incident has yet occurred. As one interviewee explained:
I don't think that the program thinks it is a very high priority to streamline any of the [data access oversight] process. I think that it will either take something bad happening and then realizing that we need additional capacities on [DAC], or some other hiccup to really promote that need.
Participant O.
Since funding for data governance is not consistently incorporated into research grants, investigators may have limited incentive to undertake the extra, often uncompensated, responsibilities associated with data governance. Inadequate allocation of resources for data sharing and governance during the planning of research studies can create obstacles for the effective implementation of data governance once the research data are deposited, as noted by one DAC member we interviewed:
We found that some people don't prioritize [data governance] because it's not helpful to them, because it's not our primary function as a department. You know, we're producing new data. That's usually what people, researchers are doing. They're not thinking about what happens to their old data. So, it's not much of a priority. Having said that, research funders are getting very keen for us to use their data. So, there is that sort of tug [of war]. … If I go into a senior team meeting, you know, something else will be the priority.
Participant F.

Structural characteristics of an organization

Our analysis revealed that several structural attributes of an institution—such as its age, staffing levels, and database size—were closely linked to perceived barriers in implementing ADS. Many participants were affiliated with DACs established within the past 1–3 years, coinciding with the creation of their institution's database. As datasets expand and attract more researchers, the risk of overwhelming existing management processes increases. It is precisely during

this formative phase that DACs could benefit from proactively considering ADS adoption and identifying potential barriers before implementation. Some participants preferred to gain experience with current data access workflows before integrating ADS, as illustrated by one member: "*because we're not sure how [name of participant's country] citizens feel or consider about the automatic decision on data sharing." – Participant K*.

Cost

Although financial constraints were less of a concern at well-funded, large-scale repositories, they posed a notable barrier for DAC members at smaller repositories, individual research departments, or genomics consortia relying primarily on grants or contracts rather than dedicated institutional funding. One participant explained:

"We [data governance office] are supported through project-specific funding. ... Governance ends up being a little bit of this indirectly supported component of our work and services. Rat has limited the ways in which we can innovate around governance. ... We don't have a huge budget."

– Participant N.

Without dedicated funding for staff, infrastructure, and materials, participants expressed concern that initial investment in ADS, substantial workflow changes, creation of training materials, and updates to internal policies would pose significant challenges.

Lack of human oversight

While some participants welcomed the potential for ADS to enhance efficiency and consistency, all emphasized that fully automating access management was unacceptable. As one interviewee noted: "*no matter what we do with automation that I feel there always needs to be that human element who's coming in and checking. So, there will always be that barrier to upscaling" – Participant E*. Others stressed the importance of understanding how research participants and the broader public would perceive ADS before implementing it. Participants were also skeptical that automated systems could adequately navigate complex, sensitive issues surrounding data reuse, which require nuanced understanding of ethical, legal, and sociocultural contexts. Several DAC members reported asking data requesters to clarify study objectives and justify their dataset needs, recognizing the importance of these sociocultural considerations.

I'm also someone who thinks that it's important to be very critical about what's the nature of the work being done. Maybe it's solid from a scientific point of view. But are there other concerns from other perspectives that need to be taken into account? That is partly why we have com- munity members on the [committee], and that's something I'm not sure can be simplified or automated."

However, when it comes to automating anything that requires reviewing information where there might be a lot of nuances, where there might be a lot of interpretation that's required, I'm a little bit more hesitant simply because I think to some extent you do need some room for a little bit of mulling over the information, ... and I think there are some information that come through with requests, that don't neatly fit into check boxes.

Participant B.

Discussion

Participants generally saw ADS tools as promising aids for DACs in enhancing efficiency and ensuring compliance in genomic data access management. Although regarded as potentially useful, these tools were not expected to directly advance the primary objectives of research organizations, such as generating high-quality data or driving scientific innovation. A major barrier identified was the low institutional prioritization of adaptable data infrastructures, alongside insufficient investment to support systems that could evolve with the increasing scale and complexity of genomic data. Many participants described ADS and related workflow improvements as secondary concerns, subordinate to regulatory compliance, investigator support, and database curation, which already consume significant DAC member time.

Funding mechanisms compound this challenge. While research grants often cover data collection and analysis, they rarely provide for the creation of governance frameworks required for effective access management. In repositories where leadership actively endorsed ADS, participants reported smoother adoption, whereas its absence—especially in smaller institutions or research labs—hampered uptake. Convincing institutional leaders

of the tangible benefits and added value of investing in data governance infrastructures emerged as a key step for raising ADS adoption as a higher priority.

Delays in updating data management infrastructure may hinder long-term repository utility. Several participants suggested that researchers are drawn to datasets primarily because of quality and diversity rather than streamlined access processes. However, evidence from genetic research indicates that ease of access, while not the primary factor, can influence researchers' database selection [28]. Repositories that proactively invest in efficient, scalable, and compliant access decision processes may therefore attract more users than those that maintain outdated workflows. Funders also influence the pace of adoption, as researchers face growing expectations to maximize output under constrained resources and tight deadlines.

Streamlining workflows was consistently highlighted as a key benefit of ADS. Participants were most enthusiastic about automating repetitive, time-intensive tasks, such as initial review and quality control of data access request forms, which are prerequisites for initiating DAC decisions. By automating these tasks, ADS could free up DAC members to concentrate on ethical deliberations and more complex decision-making.

Historically, data governance has been perceived as a secondary activity. However, new research and federal policies, such as the National Institutes of Health Data Management and Sharing (DMS) Policy, have increased its prominence by imposing additional data sharing obligations [29]. The DMS Policy exemplifies the legislative and institutional pressures that shape DAC practices, governance structures, and adoption of tools like ADS. These broader contexts informed participants' perspectives and influenced factors critical to implementation, including "structural characteristics of the organization."

The DMS Policy is poised to trigger a dramatic increase in the volume of datasets. In the absence of supportive interventions, such as ADS, institutions may face sharply rising costs for both data storage and management. Our participants indicated that many databases and repositories are created primarily to share research data funded by federal grants, often without considering existing repositories or infrastructures for data deposition. While these "blind" database initiatives are usually motivated by good intentions, they can unintentionally generate multiple, fragmented access points, leaving data technically "shared" yet difficult to locate. This is an area where ADS tools could play a critical role. One participant illustrated the challenge by describing the transfer of legacy data from a repository facing permanent closure, highlighting the need for solutions that ensure efficient and sustainable management and sharing of data even if the originating repository ceases operation. Moreover, participants noted the importance of having contingency plans for publicly funded data held in non-publicly supported repositories, in case of closures, policy changes, or staff turnover. Standardized ADS systems could facilitate interoperability between different types of repositories and enable smooth transfer of legacy data when necessary.

Limitations

Several methodological factors limit the interpretation of our results. While our participants represented diverse geographic regions, many were affiliated with DACs at large, well-resourced research institutions. It is plausible that implementation perspectives would differ substantially in smaller or under-resourced institutions. Additionally, our study relied on self-reported descriptions of institutional data access policies and procedures. Several participants were familiar with the Global Alliance for Genomics and Health and the data access committee review standards developed by our research team [30]. Although we endeavored to foster an open environment for candid discussion, social desirability bias related to our prior work may have influenced responses. Finally, because CFIR specifies a fixed set of sociological constructs relevant to implementation, our analysis was constrained to these predefined factors. Alternative analytic frameworks might have revealed additional insights that could have emerged inductively.

Conclusion

In this study, we presented findings from semi-structured qualitative interviews with DAC members worldwide, focusing on the barriers and facilitators involved in implementing ADS for genomic data access management. Overall, participants expressed general support for pilot studies evaluating ADS performance in specific workflow tasks, including cataloging data types, verifying user credentials, and tagging datasets according to use terms. Importantly, participants emphasized that

ADS should complement, rather than replace, DAC member work. This view was particularly pronounced for tasks requiring nuanced judgment and sensitivity, such as privacy protections, potential group harms, and assessing study purposes. Despite these cautions, our findings highlight cautious optimism that algorithms, software, and machine-readable ontologies could enhance efficiency, consistency, and fairness in DAC decision-making.

Based on these insights, we offer practical recommendations for institutional data stewards considering or already implementing ADS for data access management. First, repositories and institutions supporting databases and related resources should prioritize infrastructural upgrades and ensure these are reflected in associated budgets. Adequate investment in both human and material resources is critical to maintain repository utility as the volume and complexity of genomic and associated health datasets increase. Second, DACs should proactively establish data access management and sharing processes aligned with anticipated future needs. Early engagement with executive leadership is essential for integrating ADS or semi-automated tools into workflows. Supporting the case for automation with concrete trend data—such as the frequency of data access requests relative to decision turnaround time—can help anticipate repository demand over 1-, 5-, and 10-year horizons. Transparent tracking and reporting of request volumes, access decisions, and committee operations not only supports internal governance but also demonstrates responsible data stewardship to prospective contributors.

Third, DACs should avoid fully automating data access processes without maintaining human oversight over request intake and decisions. Pilot implementations should focus on tasks that are the most time-consuming, with careful inventory of required inputs along the workflow. Fourth, DACs need to assess the human and material resources necessary for effective ADS integration, including expertise, computing infrastructure, software development, and training for committee members. Finally, collaboration is critical in establishing consensus standards for adjudicating data access requests and tailoring ADS tools accordingly. Efforts such as the Ethical Provenance Subgroup of the Global Alliance for Genomics and Health (including the "Ethical Provenance Toolkit") exemplify initiatives to align practices across repositories; broader representation from institutions managing diverse health datasets would further support coordinated access management strategies.

The growing scale and complexity of genomic and associated health data, combined with the increasing need for efficient access management, underscores the potential value of ADS solutions to prevent bottlenecks and safeguard both the scientific and social value of publicly funded research data. Future research should evaluate how ADS impacts access decisions and outcomes by comparing institutions that do and do not employ such tools. These studies should determine whether ADS achieves its intended efficiency benefits and frees DAC members from procedural tasks, allowing greater focus on substantive ethical deliberations.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Stephens ZD, Lee SY, Faghri F, Campbell RH, Zhai C, Efron MJ, Iyer R, Schatz MC, Sinha S, Robinson GE. Big Data: Astronomical or Genomical? PLoS Biol. 2015;13:e1002195. https://doi.org/10.1371/journal.pbio.1002195.
2. IQVIA Institute for Human Data. Science understanding the Global Landscape of genomic initiatives: Progress and Promise.
3. Cook-Deegan R, McGuire AL. Moving beyond Bermuda: sharing data to build a medical information commons. Genome Res. 2017;27:897–901. https://doi. org/10.1101/gr.216911.116.
4. Evans B. Genomic data commons. Governing Privacy in Knowledge Commons; 2019. pp. 75–105.
5. Contreras JL, Knoppers BM. The genomic commons. Annu Rev Genom Hum Genet. 2018;19:429–53. https://doi.org/10.1146/ annurev-genom-083117-021552.
6. Grossman RL. Ten lessons for data sharing with a data commons. Sci Data. 2023;10:120. https://doi.org/10.1038/s41597-023-02029-x.
7. Cheah PY, Piasecki J. Data Access committees. BMC Med Ethics. 2020;21:1–8. https://doi.org/10.1186/s12910-020-0453-z.

8. Murtagh MJ, Blell MT, Butters OW, Cowley L, Dove ES, Goodman A, et al. Better governance, better access: practising responsible data sharing in the METADAC governance infrastructure. Hum Genomics. 2018;12. https://doi.org/10.1186/s40246-018-0154-6.
9. Lawson J, Rahimzadeh V, Baek J, Dove ES. Achieving Procedural Parity in Managing Access to genomic and related Health data: A Global Survey of Data Access Committee members. Biopreserv Biobank. 2023;bio20220205. https:// doi.org/10.1089/bio.2022.0205.
10. Shabani M, Thorogood A, Borry P. Who should have access to genomic data and how should they be held accountable? Perspectives of Data Access Committee members and experts. Eur J Hum Genet. 2016;24:1671–5. https:// doi.org/10.1038/ejhg.2016.111.
11. Ramos EM, Din-Lovinescu C, Bookman EB, McNeil LJ, Baker CC, Godynskiy et al. A Mechanism for Controlled Access to GWAS Data: Experience of the GAIN Data Access Committee. Am J Hum Genet. 2013;92:479–88. https://doi.org/10.1016/j.ajhg.2012.08.034.
12. Shabani M, Knoppers BM, Borry P. From the principles of genomic data sharing to the practices of data access committees. EMBO Mol Med. 2015;7:507–9.https://doi.org/10.15252/emmm.201405002.
13. Mello MM, Triantis G, Stanton R, Blumenkranz E, Studdert DM. Waiting for data: barriers to executing data use agreements. Science. 2020;367:150–2. https://doi.org/10.1126/science.aaz7028.
14. Lawson J, Cabili MN, Kerry G, Boughtwood T, Thorogood A, Alper P, et al. The Data Use Ontology to streamline responsible access to human biomedical datasets. Cell Genomics. 2021;1:100028. https://doi.org/10.1016/j.xgen.2021.100028.
15. Lawson J. Empirical validation of an Automated Approach to Data Use Over- sight. Cell Genomics Forthcoming; 2021.
16. Van der Auwera G, O'Connor B. Genomics in the Cloud: using Docker, GATK, and WDL in Terra. 1st ed. O'Reilly Media; 2020.
17. Birger C, Hanna M, Salinas E, Neff J, Saksena G, Livitz D, et al. FireCloud, a scalable cloud-based platform for collaborative genome analysis: strategies for reducing and controlling costs. (Bioinformatics). 2017. https://doi.org/10.1101/209494.
18. Rahimzadeh V, Lawson J, Rushton G, Dove ES. Leveraging algorithms to improve decision-making workflows for genomic data Access and Management. Biopreserv Biobank. 2022. https://doi.org/10.1089/bio.2022.0042.
19. Data Use Ontology. (2022).
20. Lawson J, Ghanaim EM, Baek J, Lee H, Rehm HL. Aligning NIH's existing data use restrictions to the GA4GH DUO standard. Cell Genomics. 2023;3:100381. https://doi.org/10.1016/j.xgen.2023.100381.
21. Rahimzadeh V, Lawson J, Baek J, Dove ES. Automating justice: An ethical responsibility of computational Bioethics. Am J Bioeth. 2022;22:30–3. https:// doi.org/10.1080/15265161.2022.2075051.
22. Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. Science. 2019;366:447–53. https://doi.org/10.1126/science.aax2342.
23. Hartzog W, Conti G, Nelson J, Shay LA. Inefficiently Automated Law Enforcement. Mich State Law Rev. 2016;1763:1763–96.
24. Ledford H. Millions affected by racial bias in healthcare algorithm. Nature. 2019;574:608–10.
25. Damschroder LJ, Aron DC, Keith RE, Kirsh SR, Alexander JA, Lowery JC. Fostering implementation of health services research findings into practice: A consolidated framework for advancing implementation science. Implement Sci. 2009;4. https://doi.org/10.1186/1748-5908-4-50.
26. Rahimzadeh V, Serpico K, Gelinas L. Institutional review boards need new skills to review data sharing and management plans. Nat Med. 2023;1–3. https://doi.org/10.1038/s41591-023-02292-w.
27. Cabili MN, Lawson J, Saltzman A, Rushton G, O'Rourke P, Wilbanks J, et al. Empirical validation of an automated approach to data use oversight. Cell Genomics. 2021;1:100031. https://doi.org/10.1016/j.xgen.2021.100031.
28. Trinidad MG, Ryan KA, Krenz CD, Roberts JS, McGuire AL, De Vries R, et al. Extremely slow and capricious: A qualitative exploration of genetic researcher priorities in selecting shared data resources. Genet Sci. 2023;25:115–24. https://doi.org/10.1016/j. gim.2022.09.003.

29. Final NIH. Policy for Data Management and Sharing https://grants.nih.gov/ grants/guide/notice-files/NOT-OD-21-013.html.
30. Global Alliance for Genomics and Health. Data Access Committee Guiding Principles and Procedural Standards Policy; 2021.