ABI African Bioinformatics Institute

Terms of Reference of the Community of Practice on Bioinformatics for Rare Diseases in Africa (CoP-RD)

1. Purpose or Rationale

Patients with rare diseases in Africa remain largely underserved, which limits the generation of knowledge that could contribute not only to African patients themselves but also to the global rare diseases field. A Community of Practice on bioinformatics for rare diseases is therefore urgently needed to collectively address the following challenges:

First, there is a severe shortage of a dedicated workforce for rare diseases in Africa. Existing gaps span across trained physicians, nurses, genetic counsellors, laboratory technicians, laboratory scientists, and bioinformaticians with rare disease expertise. A CoP can help pool scarce expertise, create training opportunities, and build a network of professionals dedicated to responding to the needs of African patients.

Second, proper clinical assessment of African patients with rare diseases is difficult due to the limited number of clinicians and the scarcity of African-specific training material. Most training materials are developed from non-African populations, limiting physicians' ability to recognize rare diseases or detect dysmorphic features in African patients. While AI-enabled diagnostic tools hold promise, these are often trained on datasets with little African data, leading to reduced accuracy. A CoP can act as a platform to coordinate efforts that generate, share, and integrate African data into AI tools, thereby improving diagnostic accuracy for African patients.

Third, access to genomic sequencing has been extremely limited. Current testing of large cohorts mostly depends on research projects or philanthropic initiatives led by non-African institutions or commercial entities. In recent years, the international effort to support the production of African data has increased. There is a need for African leaders to engage collectively, ensuring that partnerships respect African societal values and ethics. A CoP provides the collective voice and structure required for such engagement, strengthening Africa's position in global collaborations.

Fourth, Next Generation Sequencing (NGS) has become the standard test for rare diseases. This test produces a large amount of data requiring a large storage space and a high computation capacity. The lack of a big data analysis infrastructure severely restricts the continent’s ability to analyse data locally. While cloud-based storage and computation solutions offer an alternative, most infrastructures are outside Africa and governed by foreign laws, and local expertise to use emerging African-based solutions is scarce. A CoP can play a pivotal role in building local capacity, sharing expertise, and creating pathways for sustainable and continent-based data analysis.

Fifth, AI is increasingly being integrated into NGS data analysis, reducing the need for advanced computational skills and accelerating results. However, without organized African participation, these tools risk being developed without considering the continent’s realities. A CoP can help African scientists and clinicians collectively evaluate, adapt, and even co-develop AI-enabled solutions that are relevant and beneficial for African patients.

Finally, the clinical interpretation of NGS results relies heavily on public genomic databases such as gnomAD and ClinVar, which currently underrepresent African data. This leads to a disproportionate number of variants in African patients being classified as “variants of uncertain significance,” delaying or preventing diagnosis. A CoP can coordinate efforts to generate, curate, and contribute African genomic data into these global repositories, thereby improving diagnosis and contributing to global equity in rare disease research.

In sum, establishing a Community of Practice will allow African professionals to collaborate, build expertise, and strengthen leadership in bioinformatics for rare diseases, ensuring that African patients are no longer left behind in the global rare disease field.

Aims of the Community of Practice:

1) Initiate and support the development of bioinformatics capacity by training and mentoring clinicians, genetic counsellors, laboratory scientists, and bioinformaticians in genomic data analysis specifically applied to rare diseases in Africa.
2) Enhance bioinformatics support for the rare diseases community in Africa by creating African-centric digital resources, integrating African datasets into AI-enabled diagnostic pipelines, and adapting algorithms for improved recognition of rare diseases in African patients.
3) Facilitate access to sequencing and downstream bioinformatics analysis by establishing frameworks that allow African-led oversight of collaborations, ensuring sequencing data are processed and interpreted within ethical, socially relevant, and technically sound bioinformatics environments.

4) Expand local bioinformatics infrastructure for rare diseases data analysis by promoting the use of cloud and hybrid computational solutions, strengthening in-continent capacity, and building networks for resource and data-sharing across African institutions.
5) Advance the use of AI in bioinformatics pipelines by critically assessing existing tools, adapting them for African genomic diversity, and co-developing AI-enabled solutions that accelerate and simplify rare disease data interpretation.
6) Improve variant interpretation through African bioinformatics contributions by generating, curating, and depositing African genomic and clinical data into global databases, thereby enhancing the accuracy and utility of variant classification for African patients.

2. Focus of the CoP on Rare Diseases

The Community of Practice (CoP) on Rare Diseases will focus on strengthening bioinformatics capacity, infrastructure, and data resources for rare diseases in Africa. Its focus areas and examples of bioinformatics-driven activities include, without being limited to :

1) *Build bioinformatics expertise for rare diseases*
    a. Deliver training workshops on bioinformatics methods for NGS data analysis, variant annotation, and multi-omics integration for rare diseases.
    b. Establish mentorship programs connecting African bioinformaticians with junior scientists, clinicians, and students interested in rare disease genomics.
    c. Develop open-access bioinformatics training modules and case studies tailored to African data and patients.
2) *Promote African-relevant bioinformatics solutions for diagnosis*
    a. Generate and curate African genomic and phenotypic datasets to enrich bioinformatics tools used for variant prioritization and clinical decision-making.
    b. Collaborate with AI-enabled bioinformatics platform developers to retrain algorithms using African data.
    c. Build an African repository of bioinformatics-informed case studies that can support diagnostic practice.
3) *Foster equitable access to sequencing and bioinformatics analysis*
    a. Define guidelines for the ethical use and analysis of sequencing data, ensuring that African-led bioinformatics groups maintain control and oversight.
    b. Support projects that integrate local sequencing efforts with African-led bioinformatics pipelines.
    c. Advocate for African funding mechanisms to support sequencing data processing and analysis within the continent.

4) *Support bioinformatics infrastructure for big data and NGS analysis*
    a. Promote the establishment and use of African-based cloud and high-performance computing platforms for bioinformatics workflows.
    b. Contribute to the creation and sharing of bioinformatics resources that provide standardized pipelines and data storage solutions for institutions with limited infrastructure.
    c. Map and document bioinformatics infrastructure available across Africa to facilitate collaboration and resource-sharing.
5) *Advance the use of AI in bioinformatics for rare diseases*
    a. Pilot AI-based bioinformatics tools for variant prioritization, gene discovery, and phenotype-genotype correlation in African cohorts.
    b. Co-develop AI bioinformatics pipelines trained on African genomic diversity.
    c. Provide platforms for dialogue between AI developers, bioinformaticians, and clinicians to align bioinformatics tools with rare disease diagnostic needs.
6) *Contribute African data to global bioinformatics resources*
    a. Coordinate African efforts to submit curated genomic variants and clinical data into bioinformatics repositories such as gnomAD, ClinVar, and DECIPHER.
    b. Develop shared bioinformatics standards and protocols for variant curation to ensure high-quality data contribution.
    c. Publish African-led bioinformatics studies demonstrating the value of African genomic diversity in improving rare disease interpretation globally.

3. Membership and Participation

Membership is voluntary and open to individuals and institutions with relevant expertise or interest in the group’s focus area. The following categories are recognized:

1) Full Members: are actively involved in rare diseases, genomics, or bioinformatics (e.g., clinicians, researchers, bioinformaticians, data scientists, etc.). This category will include academics, practitioners, trainers, students (from undergraduate to postdoc), participating in their individual capacity or as delegates of African institutions
2) Associate Members: are individuals or organizations outside Africa who support or collaborate on projects aligned with the CoP’s vision. This category will include academics, practitioners, trainers, students, and may include representatives of institutions. Associate members are external experts who may be invited to contribute as needed.

The CoP-RD will meet monthly virtually. The CoP may also convene in-person or hybrid for a meeting or an activity when opportunity and resources permit.

While participation is voluntary, regular attendance is important. Attendance in less than 50% of activities and meetings in a year may prompt a discussion regarding ongoing involvement and capacity to contribute.

4. Roles and Responsibilities

1. *Chair/Co-Chairs*: will
    a) Facilitate meetings and discussions.
    b) Set agendas and track progress toward objectives.
    c) Represent the CoP to any relevant oversight committees or partners.
2. *Members:* will
    a) Contribute actively to discussions and deliverables.
    b) Volunteer for tasks and participate in assigned working groups.
    c) Communicate availability and progress on tasks.
3. *Secretariat (Optional Role)*
    a. Support administrative needs such as scheduling, minute-taking, and follow-ups.
    b. Maintain documentation and disseminate key materials.
    c. May be filled voluntarily on a rotational basis (e.g., every 6 months).

5. Governance and Decision-Making

Strategic and operational decisions will be taken during CoP-RD meetings. A meeting is valid when a simple quorum of at least half of the full members is attending and/or excused. The decision will be voted on by a simple majority of the members attending and/or excused.

Conflicts and disagreements will first be handled amicably within the CoP. If not resolved within the CoP, the general provision of the African Bioinformatics Institute will be applied.

6. Expected Outcomes and Timeline

6.1. *Short-Term (1–2 years)*

- Workforce and Training
    - Establishment of the Community of Practice with active membership across Africa.
    - First series of training workshops and webinars on bioinformatics for NGS and variant interpretation in rare diseases.

    - Development of initial African-focused bioinformatics training materials (case studies, e-learning modules).
- Infrastructure and Collaboration
    - Mapping of existing bioinformatics expertise, infrastructure, and gaps across African institutions.
    - Launch of small-scale shared bioinformatics pipelines accessible to member institutions.
- Ethics and Governance
    - Draft of African-led guidelines for ethical data sharing, governance, and engagement with external sequencing partners.

6.2. *Medium-Term (3–5 years)*

- Workforce and Training
    - A growing pool of trained bioinformaticians, clinicians, and scientists with practical experience in rare disease data analysis.
    - Established mentorship programs linking senior and junior African bioinformatics specialists.
- AI and Bioinformatics Solutions
    - Pilot projects demonstrating AI-enabled variant prioritization and clinical recognition tools adapted with African datasets.
    - Contribution of African phenotypic data into training sets for bioinformatics-driven diagnostic tools.
- Infrastructure and Collaboration
    - Operational shared bioinformatics hubs providing standardized NGS analysis pipelines across multiple African regions
    - Expanded access to African-based or Africa-compliant cloud platforms for genomic data storage and analysis.
- Ethics and Governance
    - Adoption of African-led ethical frameworks for sequencing and bioinformatics analysis by collaborating institutions.

7. Review and revision

These ToR will be reviewed and updated every 6 months or whether this is needed.