AZEBRA (Almost Zero Error Basepair-based Record Alert): A genomic clinical decision support system

ANAND KULANTHAIVEL, M.I.S., [1], MADHURA M. KSHIRSAGAR, M.S., [1]
MOHAMMAD ALARIFI, M.S. [2] , MARK N. OKI .AK, M.S. [1], JOSETTE F. JONES, R.N., Ph.D. [1]

[1] INDIANA UNIVERSITY, INDIANAPOLIS, IN USA --- [2] UNIVERSITY OF WISCONSIN, MILWAUKEE, WI USA

SOIC @IUPUI

Introduction

The idea of the United States's Precision Medicine Initiative (PMI) was to allow providers (and patients) to leverage large amounts of information (including patient genomic data) in order to create actionable knowledge that increases patient well-being. To this end, we propose a system called AZEBRA; the acronym stands for Almost Zero Error Basepair-based Record Alerts. Zebra, in addition to being a well-known wild animal, is a common medical slang term for the clinician's fallacy of mistakenly coming to a rare and sometimes dire diagnosis (the rare zebra diagnosis) due to having missed more common causes of patient symptoms (the common horse diagnosis); conversely, patients with rare conditions would be better thought of as zebras and not horses. AZEBRA is intended to leverage the principles of genetically-enhanced precision medicine in order to alert clinicians to the presence of patients with five (four rare, one common) genetic pathologies that are ordinarily sources of unnecessary morbidity and mortality in clinical settings.

Covered Pathologies

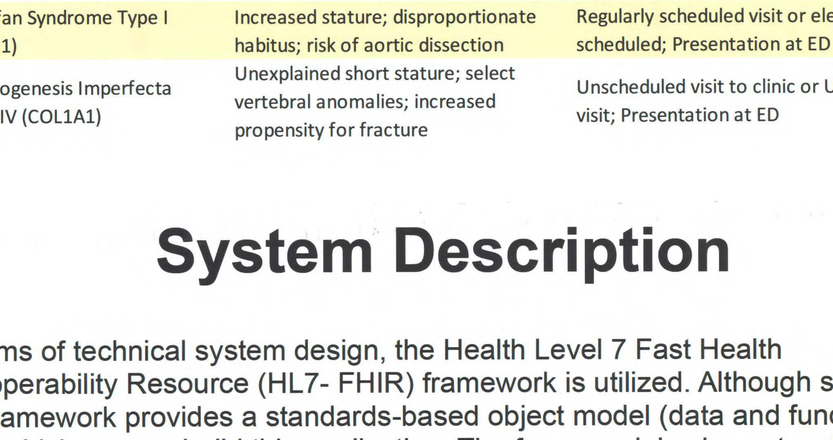
AZEBRA currently contains content to cover five genome-correlated pathological phenotypes that are encoded by 695 SNP variant mutations on 3 genes. Four of the five covered phenotypes are extremely rare (affecting less than 1/1,500 Americans). Below are descriptions of the pathologies along with scenarios (invocations) where they may be of concern.

Pathology Name & Gene	Phenotypic Aspects	Invocation situation(s)
Acromicric Dysplasia (FBN1)	Unexplained short stature; select skeletal limb anomalies	Regularly scheduled visit or elective surgery scheduled
CYP2D6 Deficiency (CYP2D6)	Increased toxicity of many commonly prescribed drugs	CPOE (Prescription) Entry
Ehlers Danlos Syndrome II/VII (COL1A1)	Elastic skin; poor healing	Regularly scheduled visit or elective surgery scheduled; unscheduled visit to clinic or urgent care clinic; EMT first response
Marfan Syndrome Type I (FBN1)	Increased stature; disproportionate habitus; risk of aortic dissection	Regularly scheduled visit or elective surgery scheduled; Presentation at ED
Osteogenesis Imperfecta I/III/IV (COL1A1)	Unexplained short stature; select vertebral anomalies; increased propensity for fracture	Unscheduled visit to clinic or Urgent care visit; Presentation at ED

System Description

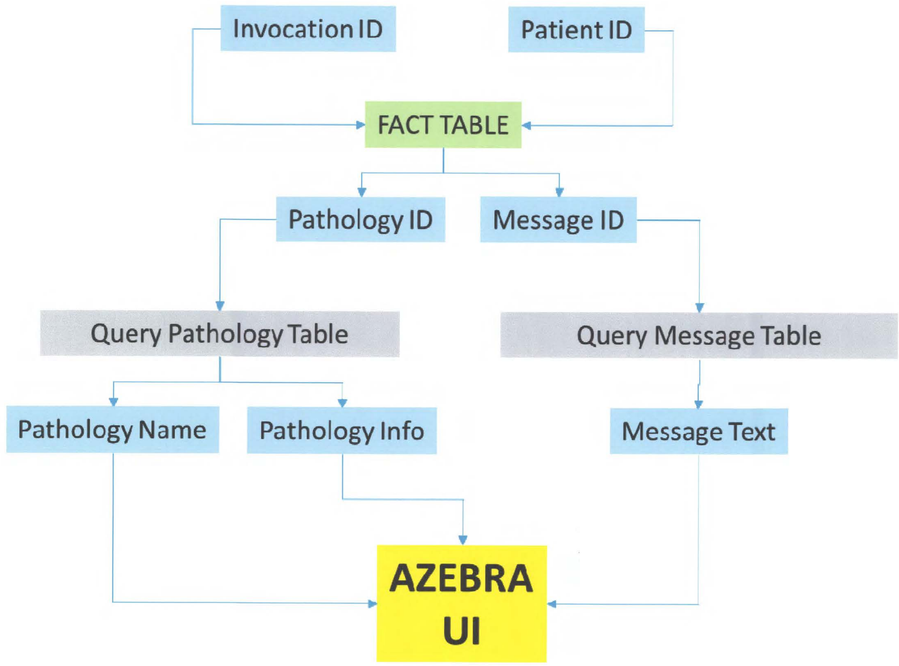
In terms of technical system design, the Health Level 7 Fast Health Interoperability Resource (HL7- FHIR) framework is utilized. Although still young, this framework provides a standards-based object model (data and functions) upon which we can build this application. The framework is also extensible, so we can use what is already available and add only the unique features (in particular, genomic data) we need for AZEBRA, thus saving software development time. Most importantly, however, AZEBRA's use of FHIR will allow it to be compatible with HL7's Version 3 format, which is increasingly forming the basis for existing enterprise EMR systems.

Figure #1: AZEBRA Data Model



For the modeling and storage of the data, we believe that a star-schema data warehouse model is most appropriate for AZEBRA due to our expectations of high query volume combined with relatively low input rates: New variant information is expected to be added only: (1) When a patient's genomic tests are complete; and (2) When new clinical intelligence supports the addition of new variants and pathologies into the system. The multidimensional analysis of the data, furthermore, lends itself to the star-schema data warehouse model in that this model can quickly accept and return not just high volumes of data, but also high complexity data.

Figure #2: Data model with interacting elements



Launch Considerations

Because existing infrastructures for the detection and storage of patient genomic data are largely inadequate for the purposes of AZEBRA, changes within the clinical workflow itself are suggested in order for a system such as ZEBRA to even be possible.

These changes include:
- Widespread sequencing of PGD (via SNP analysis)
- Patient education
- Clinician education

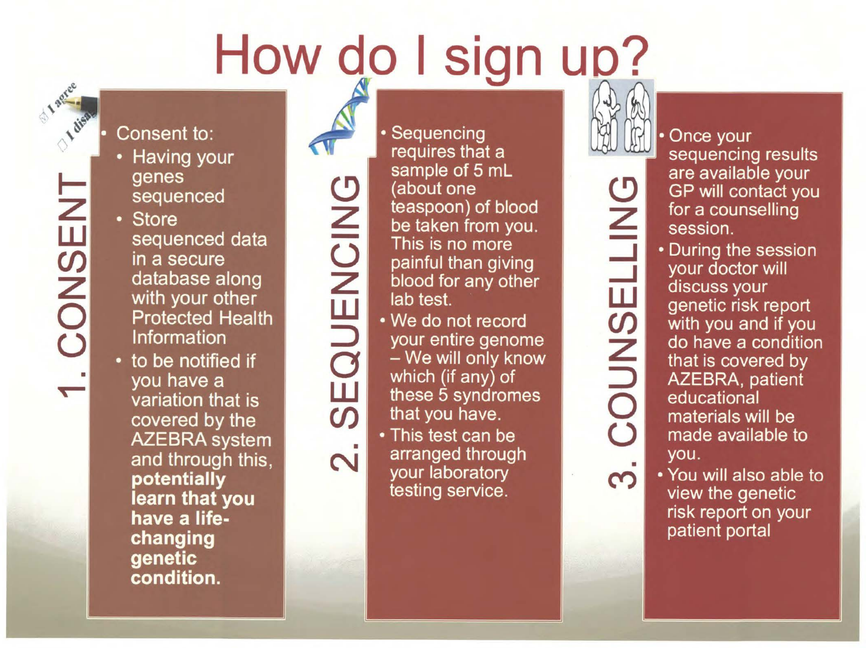


Figure #3: Patient educational material sample.

Conclusions

It is expected that the resulting system, if implemented, will reduce mortality and morbidity due to undetected pathological genomic variants. As such, AZEBRA enhances the actionability of genomic information in a meaningful fashion.

Future work:
- The actual construction, integration and loading of the required information model
- User interface design
- Usability testing of the interface and logic of the system with volunteer clinicians
  Addition of further biomedical knowledge to the system's logic

Bibliography

1. Sotos, J., Zebra Cards: An Aid to Obscure Diagnoses. 2006, Mt. Vernon, VA: Mt. Vernon Book Systems.
2. Karaa, A., Stoler, JM, Ehlers-Danlos syndrome: An unusual presentation you need to know about. Case Reports in Pediatrics, 2013(764659).
3. Malfait, F., Wnstrump, R, DePaepa, A, Ehlers-Danlos Syndrome, Classical Type. 2013.
4. Chen, H., Descartes, M Genetics of Marfan Syndrome. eMedicine, 2014.
5. Shapiro, J., Byers, PH, Glorieux, FH, Sponseller, P, Osteogensis Imperfecta: A Translational Approach to Brittle Bone Disease. 2013: Academic Press.
6. National Library of Medicine Acromicric Dysplasia (Web Resource). 2016.
7. Ingelman-Sundberg, M., Genetic polymorphisms of cytocrome P450 2D6 (CYP2D6). Clinical consequences, evolutionary aspects, and functional diversity. The Pharmacogenomics Journal, 2005. 2005(5).