mendelian

MendelApp Virtual Gene Panel Comparison

An evaluation of the accuracy of MendelApp in comparison to a panel of expert clinicians and other gene-suggestion tools.

Background

Whole Genome Sequencing (WGS) has made big improvements to diagnosis in paediatrics[1,2] and various health systems are using it in the diagnosis and stratification of rare genetic disorders and cancers[3]. Despite its successes, WGS still fails to diagnose more than 50% of patients[1,2,3.] As we discover better ways to integrate sequencing into clinical care, we expect diagnostic yield and patient management to improve.

Last year, a study from the Melbourne Genomics Health Alliance (MGHA), where 49 patients were diagnosed, showed that combining clinical interpretation with variant prioritisation improved sequence analysis[4]. A multidisciplinary team of 11 specialist clinicians were called on to generate gene panels for patients based on their phenotypic patterns. In 40 of these 49 cases (82%) the pathogenic mutation was identified in one of the genes put forward by the clinicians. The study went on to compare two automated gene suggestions tools; Phenomizer and PhenoTips, both of which showed lower success rates (37% and 28% respectively) when the top results were evaluated.

The MendelApp was built to allow clinicians and researchers to access the predictive power of Mendelian's core technology, which currently includes results from a database of 7,000 rare diseases. Clinical signs and symptoms are entered as search terms, as a result a custom gene panel, prioritised according to the most likely causative gene, is created in seconds. These results inform decisions on further clinical tests, targeted sequencing or improve downstream variant prioritisation. Furthermore, MendelApp includes suggestive deep phenotyping functionality whereby the clinician is prompted to evaluated the patient for the most informative additional signs and symptoms. In many cases these will not have been assessed or are simply overlooked.

Related work

Many tools have been created for the analysis of genetic sequences for causative variants. Some of these, Sophia DDM and Ingenuity as examples, have found commercial success in clinic and research settings.

As variant analysis has advanced, there has been exciting work in integrating phenotypic information to improve variant prioritisation. Examples include Exomizer[5], PheVor[6] and Phenolyzer[7].

The latest clinical diagnosis processes curate candidate genes before the sequence analysis takes place, as suggested in the MGHA study[4]. In paper we compare tools that can perform this curation against a team of expert clinicians. As of today, the main solutions in the field are Mendelian, Phenomizer[8] and PhenoTips[9].

Evaluation

We ran MendelApp over the 49 WES diagnoses from the MGHA study. They have kindly provided us with the clinical features of their patients (encoded in HPO) and the responses of Phenomizer, PhenoTips and their panel of clinicians. Each response was a gene list for evaluation against the gene thought to contain a pathogenic mutation.

Panel of clinicians: this is a candidate gene list created by seven geneticists, two metabolic physicians and two pediatric neurologists within 3 weeks. These lists ranged from 1-161 genes long with a median of 23 and the diagnosis was listed in 82% of the panels.

Phenomizer: we consider the top hundred differential diagnoses for the patient's HPO terms, 37% of the diagnosis were listed.

PhenoTips: we consider the suggested diagnoses section created by PhenoTips, 28% of the diagnosis were listed.

Mendelian: as with Phenomizer we consider the top hundred diagnoses. A diagnosis is considered correct if it links to the OMIM disease or gene provided by MGHA, 71% were listed.

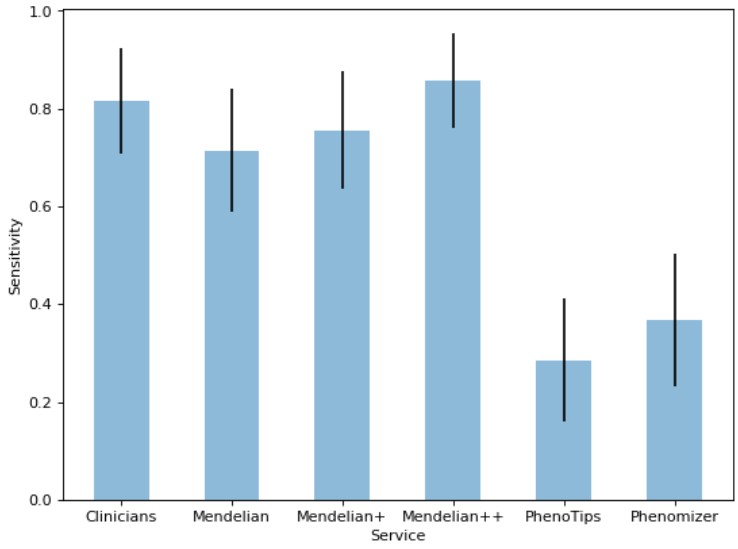
Mendelian+: as with Mendelian, except we simulate a clinician's response to the top queries generated by our deep phenotyping suggestion engine, 76% of the diagnosis were listed.

Mendelian++: Mendelian+, with the inclusion of the genes that were found in the top 161 cases (the upper limit of the clinician lists), this lifted the success rate to 86%.

Comparison of Sensitivity

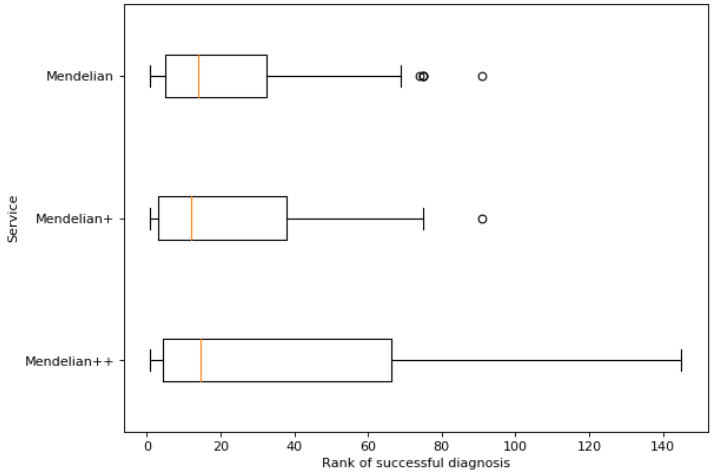
We can make a rough comparison of sensitivity by computing the number of correct predictions / number of predictions. There are a few caveats to this comparison:

- The automated tools predict diseases or disease groups, which could have none or many known genetic bases. (Mendelian tries to find the most specific diagnosis possible).
- Error bars are calculated at the 90% confidence interval



Causative Gene Prioritisation

For Mendelian we can plot the average positions causative gene in the prioritised virtual panels. The median rank of the success gene was 14 with Mendelian, 12 with Mendelian+ and 14.5 with Mendelian++.



Discussion

Mendelian's sensitivity is approaching that of expert clinicians, higher if gene panels of 161 genes are considered along with the HPO suggestion tool. This shows a marked improvement compared to the automated solutions evaluated by MGHA. The low median position of the correct gene in the priority list can be viewed as a powerful indicator of the success of the tool.

One explanation for this is the comprehensiveness of the different knowledge bases. Mendelian aggregates various databases as well as extracting knowledge from PubMed. Whereas Phenomizer is limited to OMIM, Orphanet and DECIPHER; and Phenotips to OMIM.

Another consideration is the difference in knowledge-bases between the studies. The MGHA study was conducted before this one and the growth in data during that time may give our later analysis an advantage, the clinician panels were put together in 2014 and 2015.

The gene list size is a notable differentiating factor between Mendelian and the expert clinicians. The clinicians still provide the best curation, achieving the highest sensitivity for the smallest gene lists. The longer lists generated by Mendelian my lead to higher laboratory curator burden, yet it should be noted that two factors make this difficult to assess without deeper investigation; we suspect that a longer gene list could increase the sensitivity of a diagnosis system and possibly have increased the original diagnostic yield and Mendelian's gene lists are prioritised - both of these points may reduce the laboratory burden overall.

Clinicians performed better on cases with less distinct clinical histories. For example, Mendelian occasionally missed cases that contained common terms like *Seizures*, *Hypotonia* and *Global Developmental Delay*. Each of these terms are associated to thousands of rare diseases making it more difficult to differentiate between them.

Mendelian was able to identify 4 genes missed by clinicians DYNC2H1, ECHS1D, STAT1, NALCN. Most of these cases were described with very distinct features like "*Chin with horizontal crease*" and "*Desquamation of skin soon after birth*". All of these cases were too rare to have prevalence information on Orphanet. These cases are a strong indication that a combined, clinician and Mendelian-assisted, diagnostic system could yield better results.

Conclusion

Mendelian is able to curate specific gene lists with a sensitivity close to that of a panel of expert clinicians within a matter of seconds. It has complementary strengths to the clinicians, performing better on very rare cases. This makes it a good candidate for semi-automatic curation.

Expert clinicians are experienced enough to correctly predict genes even for poorly described patients. Future work could explore emulating this in the automated systems by integrating epidemiological data, in the absence of phenotypic data.

The MGHA study mentions there were 33 patients in which no diagnosis was found. It would be interesting to see if a Mendelian-assisted curation of genes would reveal viable diagnoses for these cases.

References

1. Lionel, A.C., Costain, G., Monfared, N., Walker, S., Reuter, M.S., Hosseini, S.M., Thiruvahindrapuram, B., Merico, D., Jobling, R., Nalpathamkalam, T., Pellecchia, G., Sung, W.W.L., Wang, Z., Bikangaga, P., Boelman, C., Carter, M.T., Cordeiro, D., Cytrynbaum, C., Dell, S.D., Dhir, P., Dowling, J.J., Heon, E., Hewson, S., Hiraki, L., Inbar-Feigenberg, M., Klatt, R., Kronick, J., Laxer, R.M., Licht, C., MacDonald, H., Mercimek-Andrews, S., Mendoza-Londono, R., Piscione, T., Schneider, R., Schulze, A., Silverman, E., Siriwardena, K., Snead, O.C., Sondheimer, N., Sutherland, J., Vincent, A., Wasserman, J.D., Weksberg, R., Shuman, C., Carew, C., Szego, M.J., Hayeems, R.Z., Basran, R., Stavropoulos, D.J., Ray, P.N., Bowdin, S., Meyn, M.S., Cohn, R.D., Scherer, S.W., Marshall, C.R., 2017. Improved diagnostic yield compared with targeted gene sequencing panels suggests a role for whole-genome sequencing as a first-tier genetic test. Genetics in Medicine 20, 435–443. https://doi.org/10.1038/gim.2017.119
2. Stavropoulos, D.J., Merico, D., Jobling, R., Bowdin, S., Monfared, N., Thiruvahindrapuram, B., Nalpathamkalam, T., Pellecchia, G., Yuen, R.K.C., Szego, M.J., Hayeems, R.Z., Shaul, R.Z., Brudno, M., Girdea, M., Frey, B., Alipanahi, B., Ahmed, S., Babul-Hirji, R., Porras, R.B., Carter, M.T., Chad, L., Chaudhry, A., Chitayat, D., Doust, S.J., Cytrynbaum, C., Dupuis, L., Ejaz, R., Fishman, L., Guerin, A., Hashemi, B., Helal, M., Hewson, S., Inbar-Feigenberg, M., Kannu, P., Karp, N., Kim, R.H., Kronick, J., Liston, E., MacDonald, H., Mercimek-Mahmutoglu, S., Mendoza-Londono, R., Nasr, E., Nimmo, G., Parkinson, N., Quercia, N., Raiman, J., Roifman, M., Schulze, A., Shugar, A., Shuman, C., Sinajon, P., Siriwardena, K., Weksberg, R., Yoon, G., Carew, C., Erickson, R., Leach, R.A., Klein, R., Ray, P.N., Meyn, M.S., Scherer, S.W., Cohn, R.D., Marshall, C.R., 2016. Whole-genome sequencing expands diagnostic utility and improves clinical management in paediatric medicine. npj Genomic Medicine 1. https://doi.org/10.1038/npjgenmed.2015.12
3. Turnbull, C., Scott, R.H., Thomas, E., Jones, L., Murugaesu, N., Pretty, F.B., Halai, D., Baple, E., Craig, C., Hamblin, A., Henderson, S., Patch, C., O'Neill, A., Devereau, A., Smith, K., Martin, A.R., Sosinsky, A., McDonagh, E.M., Sultana, R., Mueller, M., Smedley, D., Toms, A., Dinh, L., Fowler, T., Bale, M., Hubbard, T., Rendon, A., Hill, S., Caulfield, M.J., 2018. The 100 000 Genomes Project: bringing whole genome sequencing to the NHS. BMJ k1687. https://doi.org/10.1136/bmj.k1687
4. Stark, Z., Dashnow, H., Lunke, S., Tan, T.Y., Yeung, A., Sadedin, S., Thorne, N., Macciocca, I., Gaff, C., Oshlack, A., White, S.M., James, P.A., 2017. A clinically driven variant prioritization framework outperforms purely computational approaches for the diagnostic analysis of singleton WES data. European Journal of Human Genetics 25, 1268–1272. https://doi.org/10.1038/ejhg.2017.123
5. Smedley, D., Jacobsen, J.O.B., Jäger, M., Köhler, S., Holtgrewe, M., Schubach, M., Siragusa, E., Zemojtel, T., Buske, O.J., Washington, N.L., Bone, W.P., Haendel, M.A., Robinson, P.N., 2015. Next-generation diagnostics and disease-gene discovery with the Exomiser. Nature Protocols 10, 2004–2015. https://doi.org/10.1038/nprot.2015.124
6. Singleton, M.V., Guthery, S.L., Voelkerding, K.V., Chen, K., Kennedy, B., Margraf, R.L., Durtschi, J., Eilbeck, K., Reese, M.G., Jorde, L.B., Huff, C.D., Yandell, M., 2014. Phevor Combines Multiple Biomedical Ontologies for Accurate Identification of Disease-Causing Alleles in Single Individuals and Small Nuclear Families. The American Journal of Human Genetics 94, 599–610. https://doi.org/10.1016/j.ajhg.2014.03.010
7. Yang, H., Robinson, P.N., Wang, K., 2015. Phenolyzer: phenotype-based prioritization of candidate genes for human diseases. Nature Methods 12, 841–843. https://doi.org/10.1038/nmeth.3484
8. Köhler, S., Schulz, M.H., Krawitz, P., Bauer, S., Dölken, S., Ott, C.E., Mundlos, C., Horn, D., Mundlos, S., Robinson, P.N., 2009. Clinical Diagnostics in Human Genetics with Semantic Similarity Searches in Ontologies. The American Journal of Human Genetics 85, 457–464. https://doi.org/10.1016/j.ajhg.2009.09.003
9. Girdea, M., Dumitriu, S., Fiume, M., Bowdin, S., Boycott, K.M., Chénier, S., Chitayat, D., Faghfoury, H., Meyn, M.S., Ray, P.N., So, J., Stavropoulos, D.J., Brudno, M., 2013. PhenoTips: Patient Phenotyping Software for Clinical and Research Use. Human Mutation 34, 1057–1065. https://doi.org/10.1002/humu.22347