

Московская школа на Юго-Западе № 1543

Кафедра биологии

Отчёт о научно-исследовательской работе

**Оценка вероятности мультифакторных  
заболеваний при помощи полногеномных скоров**

Работу выполнил  
ученик 10 "Б" класса  
А.В. Камелин

Научный руководитель:  
А.С. Ракитько

Работа выполнена в  
медико-генетическом центре  
Genotek



Москва  
2019

## Введение

Большое количество известных нам заболеваний являются мультифакторными, то есть их развитие обусловлено генетическими факторами в виде мутаций в ДНК (от точечных мутаций до хромосомных перестроек) и факторами внешней среды. Одна из важнейших задач современной науки — научиться выявлять группы риска, чтобы предсказывать эти болезни до того, как они успеют развиться. В целях изучения возможности предсказания подобных заболеваний мы решили провести исследование для такого мультифакторного заболевания, как коронарная недостаточность.

Коронарная недостаточность (КН) — патологическое состояние, характеризующееся частичным или полным прекращением коронарного кровотока и ухудшением снабжения миокарда кислородом и питательными веществами (Mayo Clinic, 2018). Основными причинами развития КН являются атеросклероз и воспаление. Увеличение атеросклеротических бляшек закрывает просвет сосудов и препятствует кровотоку, что со временем может привести к инфаркту миокарда.

Основной компонент атеросклеротической бляшки — холестерин в составе липопротеидов низкой плотности (ЛПНП), поэтому его повышенный уровень, как и уровень суммарного холестерина, является фактором риска развития атеросклероза и, как следствие, КН (Ferenec et al., 2017). Липопротеиды высокой плотности (ЛПВП), напротив, обладают антиатерогенным эффектом и удаляют избытки холестерина из тканей; их уровень отрицательно коррелирует с риском развития КН (Mahdy Ali et al., 2012). Другим важным фактором риска развития атеросклероза являются триглицериды (ТГ) — эфиры жирных кислот и глицерина (McBride, 2008; Kannel and Vasan, 2009). Они входят в состав липопротеидов очень низкой плотности и хиломикронов.

В России КН встречается у 10-14.9% населения (Zhu et al., 2016). Также КН является патогенетической основой ишемической болезни сердца, на долю которой приходится 40% всех сердечно-сосудистых заболеваний. Коронарная недостаточность обладает высокой степенью наследуемости — 50-60%

(Dai et al., 2016) — а потому представляет особый интерес для нашего исследования.

Как уже было указано ранее, на развитие КН влияет как несколько биохимических факторов (ЛПНП, ЛПВП, ТГ), так и генетические факторы. Однонуклеотидные полиморфизмы (SNP — single-nucleotide polymorphism) могут влиять на развитие самых разнообразных заболеваний: от рака щитовидной железы (Fareed and Afzal, 2013) до серповидноклеточной анемии (Shaikho et al., 2017), поскольку точечные мутации способны изменять последовательность аминокислот белка или последовательность нуклеотидов РНК, способствуя различным нарушениям и патологиям. Их исследование — полногеномный поиск ассоциаций, также известный как GWAS анализ — подразумевает поиск связи между наличием заболевания и SNP. Для этого на выборке достаточно большого размера, состоящей из здоровых и больных людей, должны сравниваться их однонуклеотидные полиморфизмы. Наиболее распространённой технологией для получения SNP в качестве данных является микрочиповое генотипирование, противопоставляемое полному секвенированию генома, требующему гораздо больших ресурсов.

Микрочиповое генотипирование — технология, позволяющая выявлять однонуклеотидные полиморфизмы. Доступность полногеномного генотипирования при помощи микрочипов привела к значительному увеличению размеров изучаемых коллекций. Технологии мета-анализов проведённых ранее полногеномных поисков ассоциаций позволяют искать генетические ассоциации на выборках из более чем 1 миллиона образцов. Результатом таких работ могут выступать полигенные скоры, основанные на оценке суммарного вклада сотен тысяч генетических маркеров. Они рассчитываются по следующей формуле:

$$PRS = \sum_{i=1}^n \beta_i x_i, \quad (1)$$

где  $\beta$  означает вес, рассчитанный для этого SNP во время GWAS анализа, а  $x$  — количество минорных аллелей.

Сам по себе микрочип (рис. 1) представляет множество ДНК-зондов — небольших нуклеотидных последовательностей, прикреплённых к некому твёрдому основанию. В основе его работы лежит явление гибридизации (комплементарные цепочки ДНК способны соединяться в двойную цепочку).

Зная, что перед тем или иным SNP у объектов нашей выборки находится консервативная нуклеотидная последовательность, мы можем сделать ДНК-зонд для каждого нужного полиморфизма, который будет комплементарен этой последовательности.



Рис. 1: Микрочип (Illumina, 2017)

Таким образом, когда мы добавим ДНК того или иного человека, определённым способом разрезанную, произойдёт гибридизация между каждым из ДНК-зондов и статичной последовательностью, предшествующей полиморфизму, в ДНК человека. Если дальше мы добавим ДНК-полимеразу и нуклеотиды, к которым прикреплена соответствующая типу нуклеотида флуоресцентная метка, то по цвету метки того нуклеотида, который присоеди-

нится к SNP, можно будет определить SNP (рис. 2).



Рис. 2: Выявление SNP методом оценки крайнего нуклеотида (Bumgarner, 2013)

Микрочипы состоят из нескольких полностью одинаковых кластеров. На одном кластере проверяется ДНК одного человека. Кластер, в свою очередь, состоит из огромного множества мелких шариков (около миллиона), в каждом из которых находится 2 ДНК-зонда для проверки диплоидного набора. Таким образом, после гибридизации и полимеризации можно будет проверить, какого цвета свечение на каждом из этих шариков при помощи специального аппарата, и мы получим данные об однонуклеотидных полиморфизмах у данного человека, что позволит нам при помощи GWAS анализа создать веса для полногеномного скора и с некоторой точностью определять вероятность развития болезни у других людей на основе их генетической информации.

Мы выбрали коронарную недостаточность в качестве исследуемого заболевания, поскольку у нас имелись данные, на которых можно было проводить исследования, а также веса полногеномного скора для этой болезни, уже полученные научным сообществом. Благодаря тому, что в данных присутствовали и генетические данные для расчёта значения полигенного скора, и характеристики внешних факторов, мы получили возможность сравнить вклад различных факторов в предсказательную способность статистических моделей, оценивающих вероятность развития данного заболевания.

## Цели и задачи

Цель: исследовать вклад генетических и негенетических факторов в методы оценки риска развития коронарной недостаточности.

Задачи:

1. Валидировать полногеномный скор, полученный в (Khera et al., 2018).
2. Выяснить, есть ли зависимость между биохимическими факторами (уровнем триглицеридов, ЛПВП и ЛПНП) и наличием заболевания.
3. Разработать несколько предсказательных моделей для коронарной недостаточности на основе данных, характеризующих генетические и негенетические факторы риска, и выявить среди них обладающую наиболее высокой предсказательной способностью.

## Материалы и методы

В качестве тестовой выборки для нашего исследования мы взяли подвыборку из PennCath study (Reilly et al., 2011), куда входило 1401 человек (из которых 933 больных и 468 здоровых) и 500 000 SNP. В качестве сопровождающей информации в ней были указаны пол, возраст, уровень триглицеридов, ЛПНП и ЛПВП.

Для исследования влияния генетических факторов на наличие заболевания мы использовали полигенный рисковый скор, основанный на GWAS анализе 6 630 150 SNP, который уже был получен в (Khera et al., 2018). Для увеличения качества выборки мы провели импьютирование (рис. 4) тестовой выборки при помощи программы BEAGLE 5.0 (Browning et al., 2018). В качестве референсной панели для импьютирования использовалась панель HRC (Consortium, the Haplotype Reference et al., 2016). Для контроля качества импьютирования применялась пост-фильтрация по  $MAF > 1\%$  и  $DR2 > 0.7$ . В результате общее количество SNP составляло 7 450 796, из которых 5 887

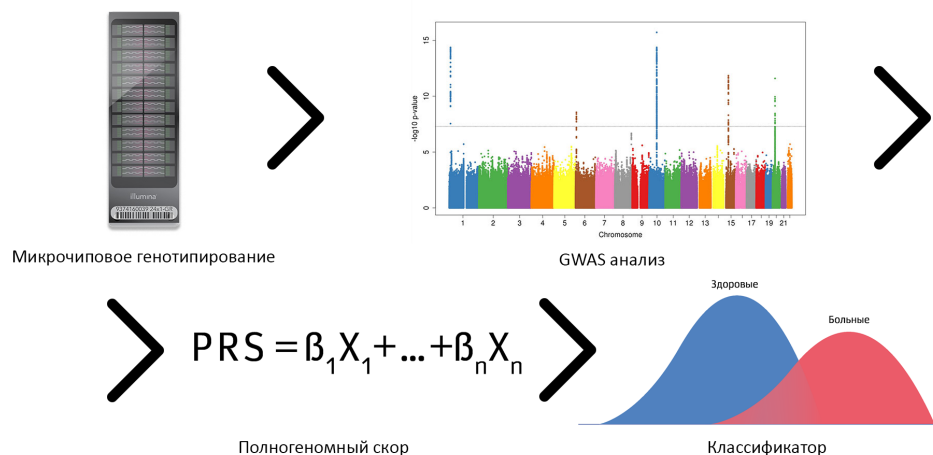


Рис. 3: Последовательность получения полногеномного скор для предсказания заболевания

713 входит в скор. Для расчёта полигенного скор для объектов выборки мы воспользовались программой PLINK v1.90b6.6 (Purcell et al., 2007; Purcell, 2007).

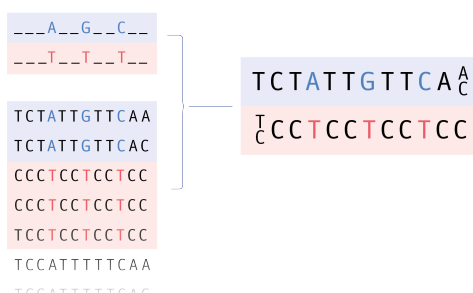


Рис. 4: Импутирование: слева сверху — обнаруженные генотипированием SNP, слева внизу — данные из референсной панели, справа — SNP после импутирования

Для визуализации результатов, построения предсказательных моделей и оценки качества предсказаний мы использовали скрипты на языке Python (Python Core Team, 2015).

Для облегчения процесса разработки мы использовали веб-оболочку для IPython под названием Jupyter Notebook (Kluyver et al., 2016) и систему управления версиями Git (Torvalds, 2005).

## Результаты

Мы построили диаграммы для трёх разных биохимических факторов (уровня триглицеридов, ЛПНП и ЛПВП), на которых были отображены распределения величин этих факторов у здоровых и больных объектов выборки.

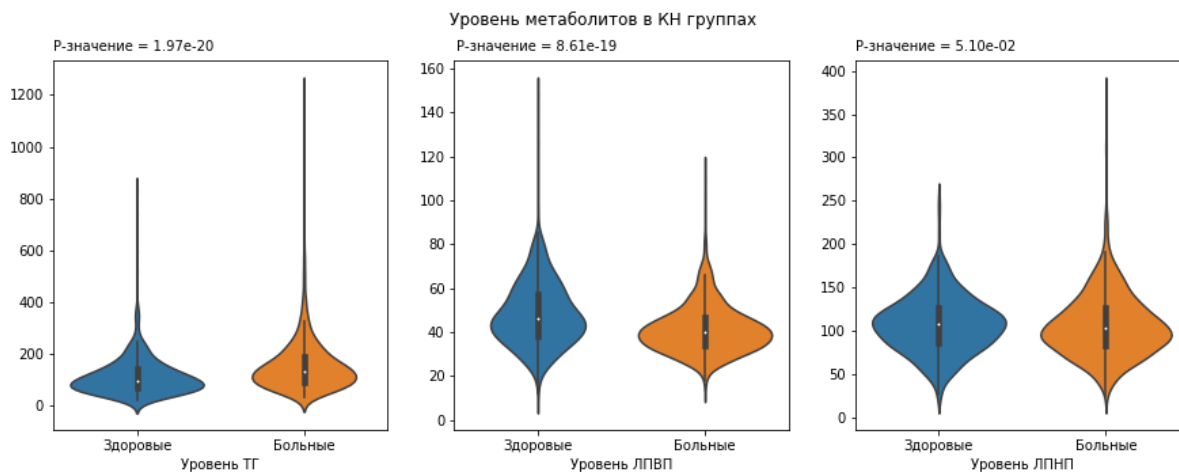


Рис. 5: Уровень метаболитов для КН групп

После этого мы провели тесты Стьюдента и Манна-Уитни для сравнения значений этих факторов в подвыборках из больных и здоровых людей. На рис. 5 в верхнем левом углу каждого подграфика указано Р-значение теста Манна-Уитни.

Из рис. 5 видно, что уровень триглицеридов у страдающих от коронарной недостаточности людей превышает уровень триглицеридов у здоровых людей. Мы обнаружили статистическую значимость этих различий (тест Стьюдента:  $p = 6.2e-12$ , тест Манна-Уитни:  $p = 1.9e-20$ ).

Соответственно, на этом графике также можно увидеть, что уровень ЛПВП у больных ниже, чем у здоровых людей. Нам удалось подтвердить статистическую значимость этих различий (тест Стьюдента:  $p = 3.1e-20$ , тест Манна-Уитни:  $p = 8.6e-19$ ).

В отличие от предыдущих биохимических факторов, для уровня ЛПНП нет чётко выраженной разницы у больных и здоровых объектов выборки. Ста-

статистические тесты также не подтвердили значимость этих различий (тест Стьюдента:  $p = 0.55$ , тест Манна-Уитни:  $p = 0.05$ ).

После мы построили несколько предсказательных моделей на основе логистической регрессии, случайного леса, градиентного бустинга и наивного байесовского классификатора. Для предсказания наличия заболевания в первом случае использовались следующие параметры: пол, биохимические факторы в виде уровня триглицеридов, ЛПНП и ЛПВП, полногеномный скор. Во втором случае обучение проводилось только на полногеномном скор, а в третьем — исключительно на негенетических факторах. Для оценки предсказательной способности мы использовали метод перекрёстной проверки.

Таблица 1: Средняя доля правильных ответов предсказательных моделей при перекрёстной проверке

Модель	Все факторы	Генетические	Внешние
Байесовский классификатор	0.762 (std: 0.03)	0.742 (std: 0.03)	0.701 (std: 0.04)
Логистическая регрессия	0.753 (std: 0.03)	0.733 (std: 0.02)	0.707 (std: 0.03)
Градиентный бустинг	0.759 (std: 0.03)	0.721 (std: 0.02)	0.708 (std: 0.04)
Случайный лес	0.745 (std: 0.02)	0.672 (std: 0.03)	0.669 (std: 0.03)

Как видно из таблицы 1, в первом случае наибольшей долей правильных ответов (accuracy) обладает наивный байесовский классификатор, за ним идёт модель на основе градиентного бустинга, логистической регрессии, а самой низкой долей правильных ответов обладает случайный лес из 1000 деревьев.

При обучении на полногеномном скор и внешних факторах средняя доля правильных ответов наивного байесовского классификатора (таблица 1) оказалась равна 0.762, а среднеквадратичное отклонение (std) — 0.03.

Во втором случае, когда обучение происходило исключительно на полногеномном скор, результаты в виде доли правильных ответов были на несколько процентов хуже, чем в первом. Наилучшую предсказательную

правильность показал также наивный байесовский классификатор (среднее: 0.742, std: 0.03).

В третьем случае доля правильных ответов была ещё хуже, чем в первом, причём в этот раз наличие или отсутствие коронарной недостаточности лучше всего предсказывала модель на основе градиентного бустинга (среднее: 0.708, std: 0.04), хотя байесовский классификатор отставал от неё на доли процентов.

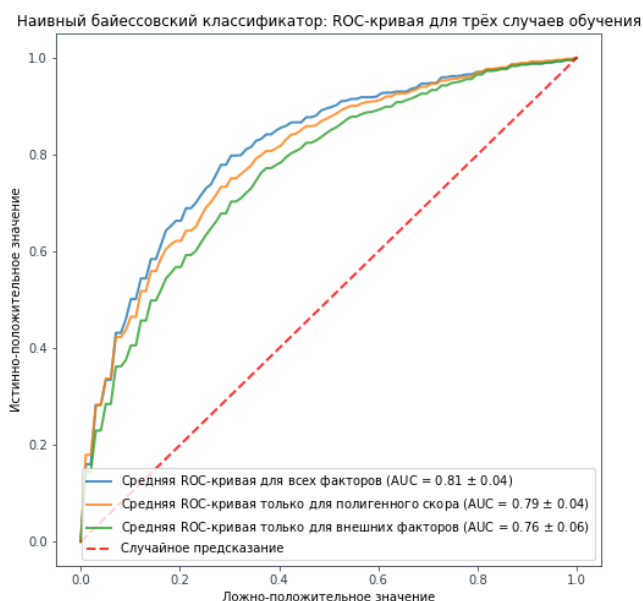


Рис. 6: Кривые ошибок наивного байесовского классификатора в трёх случаях

Если мы сравним полученные в трёх случаях  $AUC$ -показатели для модели, показавшую наилучшую правильность предсказаний — наивного байесовского классификатора, то, как видно из рис. 6, в случае использования для обучения всех данных мы получили наибольшее  $AUC$ -значение (среднее: 0.81, std: 0.04), при использовании только полногеномного сора среднее значение  $AUC$  составило 0.79 (std: 0.04), а при обучении исключительно на внешних факторах — 0.76 (std: 0.06).

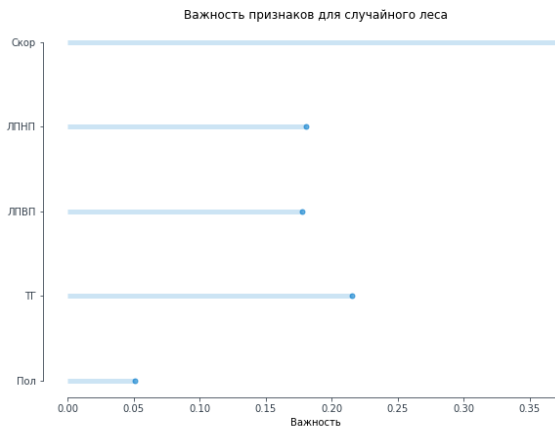


Рис. 7: Вклад отдельных факторов в предсказательную способность случайного леса

Поскольку для модели случайного леса довольно легко оценить вклад различных факторов в его предсказательную способность, мы построили диаграмму, на которой изображены величины важности всех признаков. Оказалось, что вклад полигенного сора составляет приблизительно 37% от общего, что превышает вклад каждого из внешних факторов.

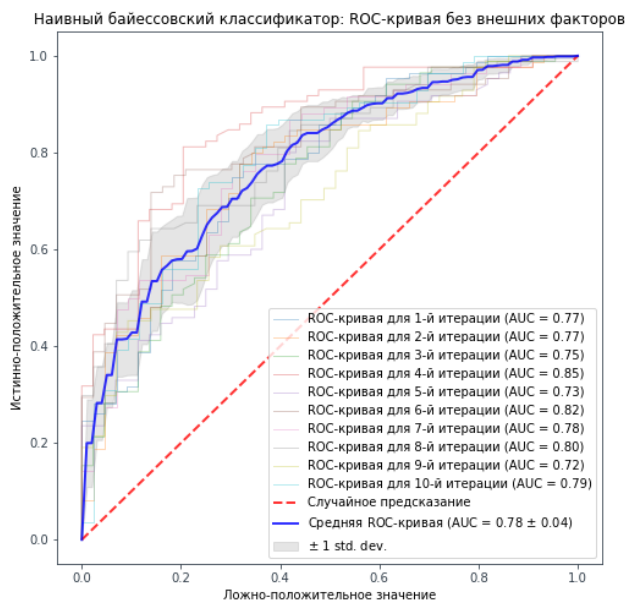


Рис. 8: Кривая ошибок наивного байесовского классификатора при обучении на полногеномном соре

По результатам видно, что обучение на генетических факторах даёт большую предсказательную способность, чем обучение на внешних факторах, а наилучшего результата удаётся достичь при комбинации и тех, и других (рис. 6). При этом важно заметить, что использование полногеномного скорра даёт довольно неплохую возможность предсказывать наличие или отсутствие коронарной недостаточности (рис. 8), а сам он вносит важный вклад в предсказательную способность (рис. 7).

## Обсуждение

Модели, обученные на генетических данных, позволяют добиться лучшей предсказательной способности, чем при обучении на внешних факторах, а значит, использование полногеномных скорров может существенно улучшить правильность предсказания возникновения мультифакторных заболеваний.

Из статистических моделей наибольшей предсказательной способностью обладает наивный байесовский классификатор, как в случае с обучением с полногеномным скорром, так и без него. В лучшем случае при обучении на всех имеющихся данных и дальнейшей перекрёстной проверке он позволяет достичь средней доли правильных предсказаний предсказаний в 0.762 (std: 0.03) и среднего AUC-показателя 0.81 (std: 0.04).

По результатам тестов мы подтвердили, что высокий уровень триглицеридов является фактором риска, способствующим развитию коронарной недостаточности, что соответствует уже имеющейся информации (Kannel and Vasan, 2009). Также, посмотрев на результаты, можно отметить, что липпротеиды высокой плотности снижают вероятность развития заболевания, что подтверждается существующими исследованиями (Mahdy Ali et al., 2012).

В отличие от других биохимических факторов, уровень липопротеидов низкой плотности не обладает статистически значимой разницей у больных и здоровых объектов выборки, что противоречит данным уже имеющихся работ, говорящих, что более высокий уровень ЛПНП является фактором риска

развития коронарной недостаточности (FERENCE et al., 2017). Возможно, это связано с неоднородностью и недостаточным размером выборки, хотя в одном из тестов P-значение было на пороге статистической значимости.

## Выводы

Использование полногеномных скоров позволяет существенно улучшить предсказательную способность статистических моделей и лучше предсказывать возникновение коронарной недостаточности. Валидирование существующих полногеномных скоров и разработка новых должны стать следующим шагом в развитии моделей предсказания мультифакторных рисков.

Повышенный уровень триглицеридов и пониженный уровень липопротеинов высокой плотности могут служить факторами риска развития коронарной недостаточности. Влияние уровня липопротеидов низкой плотности на наличие заболевания не подтвердилось в рамках этого исследования.

## Благодарности

Мы благодарим Я.В. Попова и И.И. Низамутнидова за помощь при написании работы, в получении и обработке данных. Мы также хотим выразить благодарность С.М. Глаголеву за организацию практики, директору компании Genotek В.В. Ильинскому за предоставленную возможность в проведении практики и Гвидо Ван Россуму за создание замечательного языка.

## Список литературы

Browning, B. L., Zhou, Y., and Browning, S. R. A one-penny imputed genome from next-generation reference panels. *The American Journal of Human Genetics*, 103(3):338-348, Sep 2018. ISSN 0002-9297. doi: 10.1016/j.ajhg.2018.07.015. URL <https://doi.org/10.1016/j.ajhg.2018.07.015>.

Bumgarner, R. Overview of DNA microarrays: types, applications, and their future. *Curr Protoc Mol Biol*, Chapter 22:Unit 22.1., Jan 2013.

Consortium, the Haplotype Reference, McCarthy, S., Das, S., Kretzschmar, W., Delaneau, O., Wood, A. R., Teumer, A., Kang, H. M., Fuchsberger, C., Danecek, P., Sharp, K., Luo, Y., Sidore, C., Kwong, A., Timpson, N., Koskinen, S., Vrieze, S., Scott, L. J., Zhang, H., Mahajan, A., Veldink, J., Peters, U., Pato, C., van Duijn, C. M., Gillies, C. E., Gandin, I., Mezzavilla, M., Gilly, A., Cocca, M., Traglia, M., Angius, A., Barrett, J. C., Boomsma, D., Branham, K., Breen, G., Brummett, C. M., Busonero, F., Campbell, H., Chan, A., Chen, S., Chew, E., Collins, F. S., Corbin, L. J., Smith, G. D., Dedoussis, G., Dorr, M., Farmaki, A.-E., Ferrucci, L., Forer, L., Fraser, R. M., Gabriel, S., Levy, S., Groop, L., Harrison, T., Hattersley, A., Holmen, O. L., Hveem, K., Kretzler, M., Lee, J. C., McGue, M., Meitinger, T., Melzer, D., Min, J. L., Mohlke, K. L., Vincent, J. B., Nauck, M., Nickerson, D., Palotie, A., Pato, M., Pirastu, N., McInnis, M., Richards, J. B., Sala, C., Salomaa, V., Schlessinger, D., Schoenherr, S., Slagboom, P. E., Small, K., Spector, T., Stambolian, D., Tuke, M., Tuomilehto, J., Van den Berg, L. H., Van Rheenen, W., Volker, U., Wijmenga, C., Toniolo, D., Zeggini, E., Gasparini, P., Sampson, M. G., Wilson, J. F., Frayling, T., de Bakker, P. I. W., Swertz, M. A., McCarroll, S., Kooperberg, C., Dekker, A., Altshuler, D., Willer, C., Iacono, W., Ripatti, S., Soranzo, N., Walter, K., Swaroop, A., Cucca, F., Anderson, C. A., Myers, R. M., Boehnke, M., McCarthy, M. I., Durbin, R., Abecasis, G., and Marchini, J. A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, 48:1279 EP -, Aug 2016. URL <https://doi.org/10.1038/ng.3643>.

Dai, X., Wiernek, S., Evans, J. P., and Runge, M. S. Genetics of coronary artery disease and myocardial infarction. *World J Cardiol*, 8(1):1-23, Jan 2016.

Fareed, M. and Afzal, M. Single nucleotide polymorphism in genome-wide association of human population: A tool for broad spectrum service. *Egyptian Journal of Medical Human Genetics*, 14(2):123 - 134, 2013. ISSN 1110-8630. doi: <https://doi.org/10.1016/j.ejmhg.2012.08.001>. URL <http://www.sciencedirect.com/science/article/pii/S1110863012000687>.

Ference, B. A., Ginsberg, H. N., Graham, I., Ray, K. K., Packard, C. J., Bruckert, E., Hegele, R. A., Krauss, R. M., Raal, F. J., Schunkert, H., Watts, G. F., Boren,

J., Fazio, S., Horton, J. D., Masana, L., Nicholls, S. J., Nordestgaard, B. G., van de Sluis, B., Taskinen, M. R., Tokgozoglu, L., Landmesser, U., Laufs, U., Wiklund, O., Stock, J. K., Chapman, M. J., and Catapano, A. L. Low-density lipoproteins cause atherosclerotic cardiovascular disease. 1. Evidence from genetic, epidemiologic, and clinical studies. A consensus statement from the European Atherosclerosis Society Consensus Panel. *Eur. Heart J.*, 38(32): 2459-2472, Aug 2017.

Illumina. Microarrays: Microarray analysis techniques and products, 2017. URL <https://www.illumina.com/techniques/microarrays.html>.

Kannel, W. B. and Vasan, R. S. Triglycerides as vascular risk factors: new epidemiologic insights. *Curr. Opin. Cardiol.*, 24(4):345-350, Jul 2009.

Khera, A. V., Chaffin, M., Aragam, K. G., Haas, M. E., Roselli, C., Choi, S. H., Natarajan, P., Lander, E. S., Lubitz, S. A., Ellinor, P. T., and Kathiresan, S. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nature Genetics*, 50(9):1219-1224, 2018. ISSN 1546-1718. doi: 10.1038/s41588-018-0183-z. URL <https://doi.org/10.1038/s41588-018-0183-z>.

Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., Kelley, K., Hamrick, J., Grout, J., Corlay, S., Ivanov, P., Avila, D., Abdalla, S., and Willing, C. Jupyter notebooks - a publishing format for reproducible computational workflows. In Loizides, F. and Schmidt, B., editors, *Positioning and Power in Academic Publishing: Players, Agents and Agendas*, pages 87 - 90. IOS Press, 2016.

Mahdy Ali, K., Wonnerth, A., Huber, K., and Wojta, J. Cardiovascular disease risk reduction by raising HDL cholesterol-current therapies and future opportunities. *Br. J. Pharmacol.*, 167(6):1177-1194, Nov 2012.

Mayo Clinic. Coronary artery disease - symptoms and causes, 2018. URL <https://www.mayoclinic.org/diseases-conditions/coronary-artery-disease/symptoms-causes/syc-20350613>.

McBride, P. Triglycerides and risk for coronary artery disease. *Curr Atheroscler Rep*, 10(5):386-390, Oct 2008.

- Purcell, S. *PLINK v1.90b6.6*, 2007. URL <http://pngu.mgh.harvard.edu/purcell/plink/>.
- Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., and Sham, P. C. Plink: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics*, 81(3):559-575, Sep 2007. ISSN 0002-9297. doi: 10.1086/519795. URL <https://www.ncbi.nlm.nih.gov/pubmed/17701901>. 17701901[pmid].
- Python Core Team. *Python: A dynamic, open source programming language*. Python Software Foundation, Vienna, Austria, 2015. URL <https://www.python.org/>.
- Reilly, M. P., Li, M., He, J., Ferguson, J. F., Stylianou, I. M., Mehta, N. N., Burnett, M. S., Devaney, J. M., Knouff, C. W., Thompson, J. R., Horne, B. D., Stewart, A. F., Assimes, T. L., Wild, P. S., Allayee, H., Nitschke, P. L., Patel, R. S., Martinelli, N., Girelli, D., Quyyumi, A. A., Anderson, J. L., Erdmann, J., Hall, A. S., Schunkert, H., Quertermous, T., Blankenberg, S., Hazen, S. L., Roberts, R., Kathiresan, S., Samani, N. J., Epstein, S. E., and Rader, D. J. Identification of *adamts7* as a novel locus for coronary atherosclerosis and association of *abo* with myocardial infarction in the presence of coronary atherosclerosis: two genome-wide association studies. *The Lancet*, 377(9763):383 - 392, 2011. ISSN 0140-6736. doi: [https://doi.org/10.1016/S0140-6736\(10\)61996-4](https://doi.org/10.1016/S0140-6736(10)61996-4). URL <http://www.sciencedirect.com/science/article/pii/S0140673610619964>.
- Shaikho, E. M., Farrell, J. J., Alsultan, A., Qutub, H., Al-Ali, A. K., Figueiredo, M. S., Chui, D. H. K., Farrer, L. A., Murphy, G. J., Mostoslavsky, G., Sebastiani, P., and Steinberg, M. H. A phased SNP-based classification of sickle cell anemia HBB haplotypes. *BMC Genomics*, 18(1):608, Aug 2017.
- Torvalds, L. *Git: a free and open source distributed version control system*, 2005. URL <https://git-scm.com/>.
- Zhu, K. F., Wang, Y. M., Zhu, J. Z., Zhou, Q. Y., and Wang, N. F. National prevalence of coronary heart disease and its relationship with human

development index: A systematic review. *Eur J Prev Cardiol*, 23(5):530-543, Mar 2016.