



Published in final edited form as:

*Genet Med.* 2020 September ; 22(9): 1470–1477. doi:10.1038/s41436-020-0810-9.

## Frequency of genomic incidental findings among 21,915 eMERGE network participants

Adam S. Gordon, PhD<sup>\*,1,5</sup>, Hana Zouk, PhD<sup>2,18</sup>, Eric Venner, PhD<sup>3</sup>, Christine M. Eng, MD<sup>3</sup>, Birgit H. Funke, PhD, FACMG<sup>2</sup>, Laura M. Amendola, MS, CGC<sup>5</sup>, David S. Carrell, PhD<sup>4</sup>, Rex L. Chisholm, PhD<sup>1</sup>, Wendy K. Chung, MD, PhD<sup>6</sup>, Joshua C. Denny, MD, MS<sup>16</sup>, Alexander Fedotov, PhD<sup>19</sup>, Hakon Hakonarson, MD<sup>7</sup>, Iftikhar J. Kullo, MD<sup>8</sup>, Eric B. Larson, MD, MPH<sup>4</sup>, Magalie S. Leduc, PhD<sup>3</sup>, Kathleen A. Leppig, MD<sup>19</sup>, Niall J. Lennon, PhD<sup>10</sup>, Jodell E. Linder, PhD<sup>17</sup>, Donna M. Muzny, MS<sup>3</sup>, Cynthia A. Prows, MSN, APRN<sup>11</sup>, Laura J. Rasmussen-Torvik, PhD<sup>12</sup>, Hila Milo Rasouly, PhD<sup>6</sup>, Dan M. Roden, MD<sup>15</sup>, Elisabeth A. Rosenthal, PhD<sup>5</sup>, Maureen E. Smith, MS<sup>1</sup>, Ian B. Stanaway, PhD<sup>13</sup>, Sara L. Van Driest, MD, PhD<sup>14</sup>, Kimberly Walker, MS<sup>3</sup>, Georgia L. Wiesner, MD<sup>15</sup>, Marc S. Williams, MD<sup>16</sup>, Leora Witkowski, PhD<sup>2</sup>, David R. Crosslin, PhD<sup>13</sup>, Richard A. Gibbs, PhD<sup>3</sup>, Heidi L. Rehm, PhD<sup>2,10</sup>, The eMERGE Clinical Annotation Working Group, Gail P. Jarvik, MD, PhD<sup>5</sup>

<sup>1</sup>Center for Genetic Medicine, Northwestern University, Chicago, IL

<sup>2</sup>Laboratory for Molecular Medicine, Partners Healthcare Personalized Medicine, Cambridge, MA

<sup>3</sup>Human Genome Sequencing Center, Baylor College of Medicine, Houston, TX

<sup>4</sup>Kaiser Permanente Washington Health Research Institute, Seattle, WA

<sup>5</sup>Medical Genetics, University of Washington, Seattle, WA

<sup>6</sup>Department of Pediatrics and Medicine, Columbia University Medical Center, New York, NY

<sup>7</sup>The Center for Applied Genomics, The Children's Hospital of Philadelphia, Philadelphia, PA

<sup>8</sup>Department of Cardiovascular Medicine, Mayo Clinic, Rochester, MN

<sup>9</sup>Genetic Services, Kaiser Permanente of Washington, Seattle, WA

<sup>10</sup>The Broad Institute of MIT and Harvard, Cambridge, MA

<sup>11</sup>Cincinnati Children's Hospital Medical Center, Cincinnati, OH

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:[http://www.nature.com/authors/editorial\\_policies/license.html#terms](http://www.nature.com/authors/editorial_policies/license.html#terms)

\*corresponding author. adam.gordon@northwestern.edu tel:(312)-503-1108.

Data Access: All datasets summarized here will be publicly available in the dbGaP repository under phs001616.v1.p1 and pre-dbGaP submission access can also be requested on the eMERGE website <https://emerge.mc.vanderbilt.edu/collaborate/>

**Declaration of Interests:** David Crosslin serves on a consulting board for UnitedHealth Group with precision medicine efforts, which is unrelated to this manuscript. Christine Eng is a full-time employee/faculty member of Baylor College of Medicine. Through a professional services agreement, she serves as Chief Medical Officer and Chief Quality Officer of Baylor Genetics. Richard A. Gibbs declares that Baylor College of Medicine receives payments from Baylor Genetics Laboratories, which provides services for genetic testing; Baylor College of Medicine is part owner of Codified Genomics. Heidi Rehm is employed at Massachusetts General Hospital, which receives royalties on sales of GeneInsight Software. Eric Venner is a cofounder of Codified Genomics. Georgia L. Wiesner is a member of the External Advisory Panel for the ClinGen Clinical Genome Resource Project. Hana Zouk is employed at Massachusetts General Hospital which receives royalties on sales of GeneInsight Software.

<sup>12</sup>Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

<sup>13</sup>Department of Biomedical Informatics and Medical Education, University of Washington, Seattle, WA

<sup>14</sup>Department of Pediatrics, Vanderbilt University Medical Center, Nashville, TN

<sup>15</sup>Departments of Medicine, Pharmacology, and Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN

<sup>16</sup>Genomic Medicine Institute, Geisinger, Danville, PA.

<sup>17</sup>Vanderbilt Institute for Clinical and Translational Research, Vanderbilt University Medical Center, Nashville, TN.

<sup>18</sup>Department of Pathology, Massachusetts General Hospital, Harvard Medical School, Boston, MA.

<sup>19</sup>Irving Institute for Clinical and Translational Research, Columbia University, New York, NY.

## Abstract

**Purpose**—Discovering an Incidental or Secondary Finding (IF) is a potential result of genomic testing, but little data exists describing types and frequencies of IFs likely to appear in broader clinical populations.

**Methods**—The Electronic Medical Records and Genomics network phase III (eMERGEIII) developed a CLIA-compliant sequencing panel of 109 genes and 1551 variants of clinical relevance or research interest and deployed this panel at 10 clinical sites. We evaluated medically actionable IFs across 67 genes and 14 SNVs in a diverse cohort of 21,915 participants drawn from a variety of settings (e.g. primary care, biobanks, specialty clinics).

**Results**—Correcting for testing indication, we found a 3.02% overall frequency of IFs; 2.54% from 59 genes the American College of Medical Genetics and Genomics recommends for IF return, and 0.48% in other genes, primarily *HFE* and *PALB2*. IFs associated with cancer susceptibility were most frequent (1.38%), followed by cardiovascular diseases (0.87%), and lipid disorders (0.50%). After removing *HFE*, the frequency of IFs and proportion of pathogenic vs. likely pathogenic IFs did not differ in those self-identifying as white vs. others.

**Conclusion**—Here we present frequencies and types of medically actionable incidental findings to support informed decision-making by patients, participants, and practitioners engaged in genomic medicine.

## Keywords

Personal Genomics; Clinical Sequencing; Incidental Findings; eMERGE

## INTRODUCTION

Increasing numbers of patients are undergoing panel, exome, or genome sequencing testing as a part of routine clinical care; consequently, genetic variation that is medically actionable,

but not related (incidental) to the indication for testing will become increasingly common. Though there are recommendations regarding the return of these findings<sup>1,2,21</sup>, challenges remain regarding the identification, interpretation, and return of these results to patients. To begin to address these challenges, the American College of Genetics and Genomics (ACMG) published guidelines listing 56 genes<sup>1</sup> from which incidental findings (IFs, sometimes called secondary or additional findings) should be reported when genomic sequencing is used for clinical purposes, and the types of variants, per interpretation guidelines, that would qualify as true IFs. In 2017, the ACMG working group published an updated version of these guidelines, adding 4 genes and removing one gene to make up the current list of 59 genes<sup>2</sup>.

Despite continued development and uptake of this list by the community, the most recent update to the ACMG guideline for secondary findings return still refers to pathogenicity attributes from the 2008 ACMG variant interpretation guidelines<sup>19</sup>. However, many labs return IFs classified as Likely Pathogenic (LP) or Pathogenic (P) based on the 2015 ACMG/ Association of Medical Pathology (ACMG/AMP) interpretation framework<sup>3</sup>, which is not a direct match to the Known Pathogenic (KP) and Expected Pathogenic (EP) categories recommended by current ACMG secondary findings guidelines. One major challenge to the application of the ACMG secondary findings guidelines is that, to date, much of the expert opinion regarding actionability of these gene-disease pairs does not derive from the experience of testing the broader patient populations that these guidelines are primarily intended to support. Specifically, prior estimates of IF frequency have been limited in the age, ancestry, and phenotypic diversity of the sample, restricted to a subset of all known actionable genes (e.g. hereditary cancer only), and often constrained to a single healthcare entity or testing lab<sup>9,10,16,17</sup>.

Providing data to guide the process and inform the rationale for or against IF return was a major goal of the third phase of the National Human Genome Research Institute's (NHGRI) Electronic Medical Records and Genomics network (eMERGEIII). eMERGEIII aims to study and improve standards and methods for delivery of clinical and research data across a multi-site cohort, while providing actionable genetic results derived from a next-generation sequencing platform to eMERGEIII research participants. This current phase builds on the network's decade of experience in genomics and return of results in the context of health systems, clinical data abstraction from the Electronic Health Record (EHR), linking healthcare processes that inform care with those that catalyze research, and studying multiple models of participant consent.

To this end, we developed eMERGEseq, a sequencing panel focused on 109 genes of network interest, and then selected a subset deemed 'actionable' by network consensus including all 56 genes on the initial ACMG list alongside an additional 11 genes and 14 individual variants in 11 additional genes. The panel was deployed by two CLIA sequencing labs on samples from 25,015 participants drawn from ten different clinical sites across the US<sup>4</sup>; the 21,915 participants in the IF cohort described here are participants eligible for results from the entire consensus return list who were not ascertained based on prior sequencing results. Here we describe the frequency of actionable IFs within this cohort to

better inform researchers, patients, and providers seeking an understanding of the spectrum of results that can be expected from personal genomic testing.

## MATERIALS & METHODS

### Panel development and actionability assessment:

The process to determine the eMERGEIII sequencing platform content and consensus actionability list, described elsewhere<sup>4</sup>, was led by the eMERGEIII Sequencing Centers and Clinical Annotation Working Group (ClinAnn WG), with input from the Return of Results Working Group and all eMERGE network members. At the time the eMERGEseq platform was designed, the ACMG recommended return of pathogenic variants in 56 gene-disease pairs<sup>1</sup> as incidental findings when a genomic sequencing test was ordered for a clinical indication. These genes were all included on the platform, in addition to 53 site-nominated additional genes and SNVs for inclusion on the panel, to support both discovery-focused and clinical return of result aims.

Pathogenic variants in all 56 ACMG gene-disease pairs were determined to be returnable, as well as variants in these genes pathogenic for other conditions (e.g. *WT1* NM\_024426.6:c.1447+5G>A, pathogenic for Frasier syndrome in addition to Wilms' tumor, its ACMG pair). Additional gene actionability recommendations from the literature were also considered for the remaining 53 genes; these were then individually reviewed as previously described<sup>4</sup>, considering both evidence for the gene-disease association and nature of the recommended clinical action or intervention. Additional consensus genes included *BMPRI1A*, *PALB2*, *POLD1*, *POLE*, *SMAD4*, *COL5A1*, *KCNE1*, *KCNJ2*, *HNFI1A*, *HNFI1B*, *CACNA1A*, and *OTC*; genes with individual actionable variants included *ACADM* (NM\_000016.5:c.985A>G (p.Lys329Glu)), *ALDOB* (NM\_000035.4:c.356\_359CAAA[1] (p.Asn120fs)), *BCKDHB* (NM\_183050.4:c.832G>A (p.Gly278Ser) and NM\_183050.4:c.548G>C (p.Arg183Pro)), *FAH* (NM\_000137.3:c.782C>T (p.Pro261Leu)), *G6PC* (NM\_000151.4:c.247C>T (p.Arg83Cys)), *CPT2* (NM\_000098.3:c.1239\_1240del (p.Lys414fs)), *BLM* (NM\_000057.4:c.2207\_2212delinsTAGATTC (p.Tyr736fs)), *CYP21A2* (NM\_000500.9:c.293-13C>G), *F5* (NM\_000130.4:c.1601G>A (p.Arg534Gln)), *HFE* (NM\_000410.3:c.845G>A (p.Cys282Tyr), homozygotes only), and *MEFV* (NM\_000243.2:c.2177T>C (p.Val726Ala) and NM\_000243.2:c.2080A>G (p.Met694Val))<sup>4</sup>. Though all sites participated in this review, latitude was given to individual sites regarding returning these results to account for differences in study population (e.g. pediatrics sites), study aims, consent, and institutional review.

The ACMG updated its recommendation to 59 genes in 2017, dropping 1 (*MYLK*) and adding 4 (*ATP7B*, *BMPRI1A*, *SMAD4*, and *OTC*)<sup>2</sup>. Three of these new genes were included on the eMERGEseq panel and had already been deemed actionable by the eMERGEIII network, and the fourth was represented by a single SNV (*ATP7B*; NM\_000053.4:c.3207C>A (p.His1069Gln)), which was returned only by 2 clinical sites. *MYLK* was specifically discussed and, given its calculated score per the ClinGen gene-disease validity framework<sup>22,23</sup> did not differ from other ACMG-included genes associated with thoracic aortic aneurysms and aortic dissections (TAAD), it was retained as actionable by the eMERGE network. Given the research nature of the project, the consent process, and

the participants' accessibility to medical care, after some discussion it was decided by network consensus to return both pathogenic (P) and likely pathogenic (LP) variants in all 68 genes.

### **Cohort description:**

The eMERGEIII network consists of 10 clinical study sites, two Sequencing Centers (SCs), and a Coordinating Center. A primary goal of eMERGEIII is to harmonize and implement an entire, unified pipeline simulating real-life clinical genomic testing within a broad spectrum of participants and clinical settings, with each clinical site's cohort of ~2500 participants being, by design, markedly different in ascertainment strategy and demographics. The study adhered to the principles set out in the Declaration of Helsinki, with informed consent from all participants, as approved by individual sites' IRBs; a specific discussion of the ethical considerations across the eMERGEIII study are described elsewhere<sup>25</sup> One clinical site that did not return the consensus list to participants could not be included in this analysis. Another clinical site could not be included as they ascertained participants based on previous non-CLIA sequencing results enriching for putative pathogenic variants, which would bias the analysis; results from this site have been previously described elsewhere<sup>17</sup>. The resulting IF cohort (N = 21,915) includes a diverse range of participants drawn from across the lifespan and across the health system, from primary care to specialty clinics; Table 1 details the demographic makeup of both the overall and individual site cohorts.

### **Clinical sequencing & interpretation pipeline:**

Details of the eMERGEseq clinical sequencing pipeline are described in detail elsewhere<sup>4</sup>. Briefly, participants were enrolled at each site, blood collected, DNA extracted locally, then sent to one of two assigned SCs for targeted sequencing in a CAP/CLIA setting: the Baylor College of Medicine Human Genome Sequencing Center [BCM-HGSC], Houston TX; and the Broad Institute and Partners Laboratory for Molecular Medicine [LMM], Cambridge, MA. Variants identified through this sequencing test were interpreted by the assigned SC in the context of the site-provided disease status and test indication (if present). Variant classifications from both laboratories were based primarily on ACMG/AMP criteria<sup>3</sup>, supplemented by recommendations from ClinGen's Expert Panels & Sequence Variant Interpretation Working Group. Final interpretation decisions incorporated these recommendations along with evidence from the literature and, when present, additional SC internal data accrued from previous testing. Evidence summaries from the laboratories supporting these interpretations were included on the clinical reports. Variant classifications along with evidence and interpretive summaries for all reported variants were submitted directly to ClinVar by both BCM-HGSC and LMM (submitter IDs 500199 & 21766, respectively). Though there are subtle differences between the sequencing and variant analysis approaches at the two SCs, an extensive preliminary harmonization effort between the two SCs at all critical components of the sequencing and interpretation workflow ensured that test quality and accuracy were highly concordant<sup>4</sup>.

### Working Definition & Identification of Incidental Findings:

Although there is some debate as to the specific terminology of each case, the rate of clinically actionable findings unrelated to any participant ascertainment indication can be estimated. In this study, for the purpose of clarity, an “incidental finding” (IF) is defined as a pathogenic (P) or likely pathogenic (LP) variant in a gene on the consensus list with an associated disease phenotype unrelated to any participant indication that was provided to the SCs at sample intake. For individuals without an indication for testing, any positive (P or LP) finding was considered incidental. Thus, IFs in this cohort derive from individuals both with and without a reported indication for testing. Supplementary Table 1 lists these indications and their frequency in this cohort.

Given the mix of ascertainment protocols, the distribution of those eligible to receive primary findings (i.e. those with an indication for testing) varies considerably across site and disease phenotype; this could artificially skew overall estimates of IF frequency for phenotypes that were also indications for testing. To correct for this, we adjusted the total number of participants contributing to each frequency estimate based on indication for testing. For example, in calculating the IF rate for *LDLR*, we excluded the 1,626 participants indicated to have a lipid disorder, as these participants by definition could not have received an IF in *LDLR*. Also, though a limited set of pharmacogenetic markers were also assayed and reported by the SCs, these results, known to be relatively common<sup>7</sup>, are not included in this analysis. Likewise, carrier status for conditions with autosomal recessive inheritance is not included in this analysis.

## RESULTS

### Overall IF rates:

Among the 1,166 positive reports across the entire eMERGEIII IF cohort (N=21,915), we identified 661 actionable findings unrelated to participant test indication, resulting in an overall IF rate of 3.02% (Figure 1). Of these, 556 findings (2.54%) are variants in genes recommended by ACMG to be reported as IFs, while the remaining 105 findings (0.48%) are in other consensus actionable loci. Ten individuals were found to have two independent IFs, each unrelated to participant test indication. Ten variants were each reported in 5 participants (Table 2). Three of these 10 variants, in *MYBPC3*, *KCNQ1*, and *RYR1* respectively, were reported as LP. A full list of variants determined to be IFs and their rates within this cohort can be found in Supplementary Table 2. Additionally, two participants were found to have the unanticipated IF of acquired mosaic trisomy 12, a marker in chronic lymphocytic leukemia<sup>24</sup>; after consultation between the relevant SC and clinical site, these results were included on the participant’s clinical report.

### IF rates by gene & disease domain:

IF counts by disease domain are illustrated in Figure 2a. Most IFs identified in this study were in cancer associated genes, with cardiovascular disease associated genes being the next most common group; these two domains represent the majority of genes on the IF list. For cancer associated genes, most IFs were P (1.12% P vs. 0.26% LP), while for the cardiac associated genes most were LP (0.36% P vs. 0.51% LP). Figure 2b illustrates the IFs



reported by gene. IFs were most commonly reported in *BRCA2*, *LDLR*, *HFE*, *MYBPC3*, and *BRCA1*, in that order. Of note, for *HFE*, only homozygosity for p.Cys282Tyr (NM\_001300749.2:c.845G>A) was returned.

### IF rates across self-reported ancestry groups:

Table 3 summarizes the rate of IFs by self-reported ancestry. Self-reported Caucasian/white participants had the highest rate, 3.10% (95% CI 2.84–3.40, N=14,480); followed by Asian, 2.74% (2.02 – 3.69, N=1497); Black/African-American, 2.29% (1.83 – 2.86, N=3,279); Hispanic/Latinx, 1.98% (1.41 – 2.77, N=1666); and American Indian, Alaska Native, or Native Pacific Islander, 1.30% (0.23 – 6.97, N=77). In participants where self-reported race was known, There was a significant excess of IFs in those who self-reported as White vs. those who did not when *HFE* p.Cys282Tyr (NM\_001300749.2:c.845G>A) results were included (chi-sq,p=0.003, 1df). This allele is known to be common only in European ancestry individuals<sup>8</sup> and indeed, 65/68 participants in our cohort with this finding identified as White; of the remaining 3, one identified as Black/African-American and the final 2 were unknown. After excluding this allele, frequency of IFs did not differ significantly between those self-reporting their race as White (2.68%) vs. all other groups combined (2.36%; chi-sq. p=0.20, 1df). The fraction IFs classified as LPs also did not differ between those self-reporting as White (38.1%) and other groups (41.0%; chi-square, p = 0.50).

We did not have access to data on whether participants self-reported Ashkenazi Jewish (AJ) ancestry. These individuals, representing roughly 1.73% of the US population, are known to have a 1/40 rate of harboring a pathogenic *BRCA1/2* founder allele. Based on these rates, we would expect to find ~10 individuals with one of these variants across our cohort of 21,915; we found 26 participants with one of these variants (NM\_000059.3(*BRCA2*):c.5946del (p.Ser1982fs), NM\_007294.3(*BRCA1*):c.66\_67AG[1] (p.Glu23fs)), suggesting enrichment of individuals of AJ ancestry in our cohort. This enrichment was not driven by one clinical site alone. However, excluding participants with these variants from IF analyses does not change our finding that the most frequent disease category of IFs was genes associated with cancer susceptibility (1.38%; 1.24% without AJ founder alleles), though previous studies found the highest rate to be cardiac disease associated genes<sup>16,18</sup> (0.87% in our cohort).

## DISCUSSION

As patients and research participants consider the option to receive IFs as part of a genomic test, it is important that they have information on the probability and type of IFs that may be reported. Health systems must also be prepared to return these results and offer follow-up care. To our knowledge, this report is the largest study of medically actionable incidental findings to date. In our cohort, the frequency of IFs was 3.02% overall, with 2.54% in the 59 genes recommended by ACMG to be reported as IFs. This result is consistent with previous estimates of this frequency, though prior studies differ in the genes included as IFs, precede the current ACMG/AMP variant classification system, derive from a single health system, or have considerably smaller sample size<sup>9,10,16,17,18</sup>.

Our study cohort includes more individuals identifying as non-White than most studies evaluating IF rates. After removal of the *HFE* p.Cys282Tyr variant (NM\_001300749.2:c.845G>A), we did not find a higher rate in those who self-report as White vs other race/ethnicity groups, nor did the proportion of IFs that were classified as LP vs P. This differs from some previous studies of IF rate, which had previously reported an excess of IFs in those of European ancestry vs African ancestry. Aside from our exclusion of *HFE*, this difference could also be potentially explained by their use of slightly different lists of genes to return, cohorts outside the realm of clinical genetic testing, and earlier versions of population databases with limited representation of non-White individuals<sup>9,10</sup>. While there may be differences in the rate of predicted deleterious variants among ancestry groups<sup>11,12</sup>, the frequency of medically actionable incidental findings in our cohort does not mirror this prediction, though our study was not powered to test differences among the individual groups comprising ‘non-White.’ This limitation, and the dearth of published evidence to accurately classify variants specific to other ancestry backgrounds (e.g. co-segregation), underscores the need to amplify participant recruitment across ancestry groups historically under-represented in genomics research in order to ameliorate genomic health care disparities<sup>13</sup>.

eMERGE did consider several genes to be returnable beyond those recommended by ACMG. Of these, we returned IFs in *HFE* (associated with hemochromatosis; only in homozygotes for the higher-penetrance p.Cys282Tyr variant<sup>18</sup>) *PALB2* (breast cancer), *KCNE1* (arrhythmia), *HNF1A* (maturity-onset diabetes of the young), and *MEFV* (familial Mediterranean fever; only in homozygotes or compound heterozygotes for NM\_000243.2:c.2080A>G (p.Met694Val) or NM\_000243.2:c.2177T>C (p.Val726Ala)). The p.Cys282Tyr variant in *HFE* (NM\_001300749.2:c.845G>A) was the third most common IF returned and contributed to the higher rate of IFs in those identifying as White. eMERGE has previously advocated for the return of this genotype as an IF<sup>14,8</sup>, reporting a penetrance rate of hemochromatosis of 24.4% in males and 14.0% in females with this genotype<sup>8</sup>.

The eMERGEIII network elected to return LP IFs in this research context; 43.5% of the reported results were LP. One goal of this return was to collect further data on participants who harbor an LP variant (and potentially also their family members) to assist in any potential future reclassification. In our cohort, the proportion of variants classified as LP was greater among cardiac genes than the proportion in cancer genes. Factors that may account for this difference are 1) lower or unknown penetrance for many cardiac genes; 2) a dramatically higher burden of missense variants among results (47% of cardiac IFs vs. 21% of cancer IFs in our study), which are generally classified with less certainty than putative truncations; and 3) the smaller evidence base available for variant interpretation in cardiac genes, which have generally been studied and reported clinically for a shorter time period and in fewer individuals than cancer genes<sup>20</sup>. While current ACMG/AMP guidelines do not explicitly recommend return of LP variants, they are often reported by clinical labs. We note, however, that even over the course of this study, some variants initially classified as LP (e.g. NM\_000256.3(*MYBPC3*):c.3628-41\_3628-17del25 and NM\_198056.2(*SCN5A*):c.3911C>T (p.Thr1304Met)) were downgraded by one of the SCs to VUS or below. This highlights that the true frequency of LP will likely shift over time as



new evidence emerges, new terminology for low-risk and low-penetrance loci is adopted, and the interpretation guidelines themselves continue to evolve.

Limitations of the study include smaller sample sizes for non-European ancestry groups, though this cohort is more heterogeneous than previous studies reporting IF frequencies<sup>9,10,16,17,18</sup>. Due to sample size, which limited power of inter-population comparisons, we pre-specified testing the frequency in Whites vs. all other ancestries combined. As such, differences between Whites and specific non-White groups, even if small, cannot be ruled out using these data. Despite these limitations, to our knowledge, this is the largest study of IF rates in self-identified non-Whites. Though we did not have information on self-reported AJ ancestry, removing known AJ founder alleles from the analysis does not alter our findings that IFs are most likely to be in genes associated with cancer risk, though a previous study found the highest rate to be in genes associated with cardiac disease.<sup>18</sup>

An additional limitation is that we could not independently confirm patient indication for testing via EHR review. It was difficult for a minority of cases recruited in specialty clinics to determine if a finding was truly ‘incidental.’ In some cases, participants were assigned an indication based on the specialty clinic in which they enrolled, but it is possible that some of these participants were family members of those with the condition necessitating the visit to these clinics. Regardless, ongoing detailed EHR review for participants with positive genetics results is currently underway by the eMERGE Outcomes and Clinical Annotation workgroups to determine the spectrum of clinical features in all individuals in which a P or LP variant was identified. Finally, *ATBP7*, associated with Wilson’s disease, was not fully sequenced but only genotyped for a single known P variant. However, given the rarity of that disorder (estimated as 1/30,000), our estimate of the IF rate for ACMG genes is likely not affected by its omission.

In summary, we evaluated the rate of IFs in a diverse cohort of 21,915 eMERGEIII participants. We found a 3.02% overall frequency of IFs, of which 2.54% were in the genes deemed actionable by the ACMG. The most frequent category of IFs in our cohort were in genes associated with cancer susceptibility (1.38%), followed by cardiac diseases (0.87%) and lipid disorders (0.50%). Though the frequency of IFs and proportion of P vs. LP IFs were higher in those self-identifying as White than in those identifying as another group, this difference was not significant after removing *HFE* results from the analysis. These important findings serve as resource to inform decision making in patients and research participants undergoing genomic testing, aid the ongoing development of practice standards and guidelines in genomic medicine, and drive future research efforts in variant interpretation and IF return.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

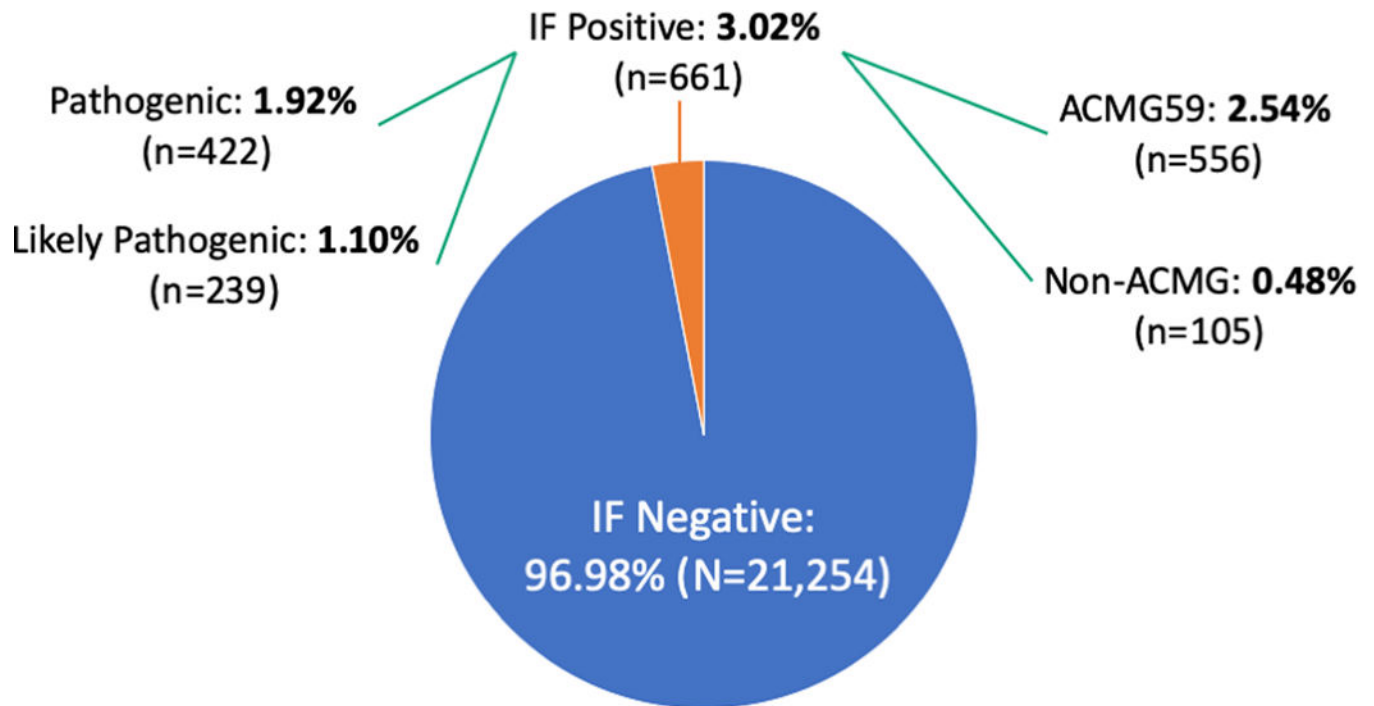
## Acknowledgements:

The eMERGE Phase III Network was initiated and funded by NHGRI through the following grants: U01HG8657 (Kaiser Permanente Washington/University of Washington); U01HG8685 (Brigham and Women's Hospital); U01HG8672 (Vanderbilt University Medical Center); U01HG8666 (Cincinnati Children's Hospital Medical Center); U01HG6379 (Mayo Clinic); U01HG8679 (Geisinger Clinic); U01HG8680 (Columbia University Health Sciences); U01HG8684 (Children's Hospital of Philadelphia); U01HG8673 (Northwestern University); MD007593 (Meharry Medical College); U01HG8701 (Vanderbilt University Medical Center serving as the Coordinating Center); U01HG8676 (Partners Healthcare/Broad Institute); and U01HG8664 (Baylor College of Medicine).

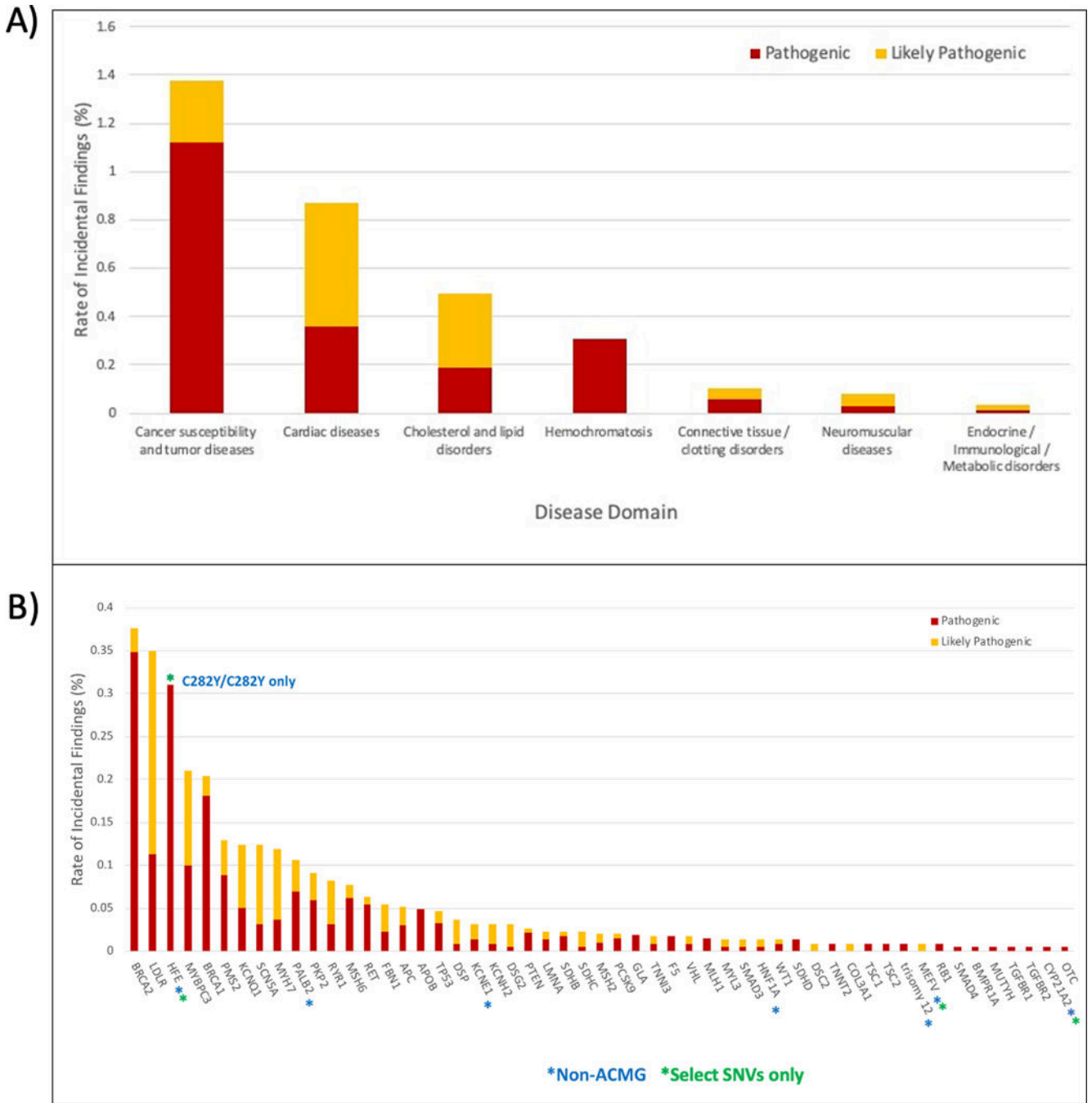
## REFERENCES

- Green RC, Berg JS, Grody WW, et al. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med*. 2013;15(7):565–74. [PubMed: 23788249]
- Kalia SS, Adelman K, Bale SJ, et al. Recommendations for reporting of secondary findings in clinical exome and genome sequencing, 2016 update (ACMG SF v2.0): a policy statement of the American College of Medical Genetics and Genomics. *Genet Med*. 2017;19(2):249–255. [PubMed: 27854360]
- Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405–24. [PubMed: 25741868]
- The eMERGE Consortium. Harmonizing clinical sequencing and interpretation for the eMERGE III network. *Am J Hum Genet*. 2019;105(3):588–605. [PubMed: 31447099]
- Dorschner MO, Amendola LM, Turner EH, et al. Actionable, pathogenic incidental findings in 1,000 participants' exomes. *Am J Hum Genet*. 2013;93(4):631–40. [PubMed: 24055113]
- Gordon A, Rosenthal E, Carrell D, et al. Rates of actionable genetic findings in individuals with colorectal cancer or polyps ascertained from a community medical setting. *Am J Hum Genet*. 2019 (in press)
- Tabor HK, Auer PL, Jamal SM, et al. Pathogenic variants for Mendelian and complex traits in exomes of 6,517 European and African Americans: implications for the return of incidental results. *Am J Hum Genet*. 2014;95(2):183–93. [PubMed: 25087612]
- Gallego CJ, Burt A, Sundaresan AS, et al. Penetrance of Hemochromatosis in HFE Genotypes Resulting in p.Cys282Tyr and p.[Cys282Tyr];[His63Asp] in the eMERGE Network. *Am J Hum Genet*. 2015;97(4):512–20. [PubMed: 26365338]
- Olfson E, Cottrell CE, Davidson NO, et al. Identification of Medically Actionable Secondary Findings in the 1000 Genomes. *PLoS ONE*. 2015;10(9):e0135193.
- Amendola LM, Dorschner MO, Robertson PD, et al. Actionable exomic incidental findings in 6503 participants: challenges of variant classification. *Genome Res*. 2015;25(3):305–15. [PubMed: 25637381]
- Lohmueller KE, Indap AR, Schmidt S, et al. Proportionally more deleterious genetic variation in European than in African populations. *Nature*. 2008;451(7181):994–7. [PubMed: 18288194]
- Fu W, O'connor TD, Jun G, et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature*. 2013;493(7431):216–20. [PubMed: 23201682]
- Fullerton SM, Knerr S, Burke W. Finding a place for genomics in health disparities research. *Public Health Genomics*. 2012;15(3–4):156–63. [PubMed: 22488458]
- Fullerton SM, Wolf WA, Brothers KB, et al. Return of individual research results from genome-wide association studies: experience of the Electronic Medical Records and Genomics (eMERGE) Network. *Genet Med*. 2012;14(4):424–31. [PubMed: 22361898]
- DellaPergola S (2013) World Jewish Population, 2012 In: Dashefsky A, Sheskin I (eds) *American Jewish Year Book 2012*. American Jewish Year Book, vol 109–112. Springer, Dordrecht.
- Yang Y, Muzny DM, Xia F, et al. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA*. 2014;312(18):1870–9. [PubMed: 25326635]

17. Schwartz MLB, McCormick CZ, Lazzeri AL, et al. A Model for Genome-First Care: Returning Secondary Genomic Findings to Participants and Their Healthcare Providers in a Large Research Cohort. *Am J Hum Genet.* 2018;103(3):328–337. [PubMed: 30100086]
18. Retterer K, Juusola J, Cho MT, et al. Clinical application of whole-exome sequencing across clinical indications. *Genet Med.* 2016;18(7):696–704. [PubMed: 26633542]
19. Richards CS, Bale S, Bellissimo DB, et al. ACMG recommendations for standards for interpretation and reporting of sequence variations: Revisions 2007. *Genet Med.* 2008;10(4):294–300. [PubMed: 18414213]
20. Harrison SM, Dolinsky JS, Knight Johnson AE, et al. Clinical laboratories collaborate to resolve differences in variant interpretations submitted to ClinVar. *Genet Med.* 2017;19(10):1096–1104. [PubMed: 28301460]
21. The use of ACMG secondary findings recommendations for general population screening: a policy statement of the American College of Medical Genetics and Genomics (ACMG). *Genet Med.* 2019;21(7):1467–1468. [PubMed: 31019278]
22. Berg JS, Foreman AK, O'daniel JM, et al. A semiquantitative metric for evaluating clinical actionability of incidental or secondary findings from genome-scale sequencing. *Genet Med.* 2016;18(5):467–75. [PubMed: 26270767]
23. Hunter JE, Irving SA, Biesecker LG, et al. A standardized, evidence-based protocol to assess clinical actionability of genetic disorders associated with genomic variation. *Genet Med.* 2016;18(12):1258–1268. [PubMed: 27124788]
24. Autore F, Strati P, Laurenti L, Ferrajoli A. Morphological, immunophenotypic, and genetic features of chronic lymphocytic leukemia with trisomy 12: a comprehensive review. *Haematologica.* 2018;103(6):931–938. [PubMed: 29748447]
25. Fossey R, Kochan D, Winkler E, et al. Ethical Considerations Related to Return of Results from Genomic Medicine Projects: The eMERGE Network (Phase III) Experience. *J Pers Med.* 2018;8(1)



**Figure 1: Overall frequency of Incidental Findings across the eMERGEIII-IF cohort.** 10 individuals were found to have 2 different IFs.



**Figure 2: Incidental Finding rates grouped by (A) associated disease domain and (B) gene.** Interpretations as reported by the sequencing centers. Rates are calculated using the proportion of participants that do not have a related indication as enumerated in Supplementary Table 1. Hemochromatosis findings are *HFE* p.C282Y homozygotes only (NM\_001300749.2:c.845G>A).

**Table 1:** Demographic characteristics of the eMERGE3 Incidental Findings cohort and 9 contributing clinical sites.

Self-reported race/ethnicity derives from multiple-choice selection of US Census-standardized terms and does not fully capture the complexity of participant racial identity. CCHMC = Cincinnati Children’s Hospital Medical Center; KPW/UW = Kaiser Permanente Washington / University of Washington; CHOP = Children’s Hospital of Philadelphia. Sequencing Centers (SCs, italics): The Baylor College of Medicine Human Genome Sequencing Center [BCM-HGSC], Houston TX; Broad Institute and Partners Laboratory for Molecular Medicine [LMM], Cambridge, MA.

eMERGE Site	N	Self-reported race/ethnicity							Median Age (yrs)	Age Range (yrs)
		Hispanic or Latinx	Black or African-American	Asian	American Indian, Alaska Native, or Pacific Islander	White	Unknown / Not Reported	% Female		
CCHMC	2956	24	1041	16	2	1697	176	34	20	4 – 85
Harvard	2492	146	145	72	0	2038	91	57	61	22 – 100
KPW/UW	2495	43	51	996	51	1297	57	61	65	32 – 104
<i>SC:LMM</i>	<i>7943</i>									
CHOP	2972	126	1171	42	3	1524	106	21	18	7 – 32
Columbia	2566	636	302	198	9	1072	349	55	55	21 – 97
Mayo	3018	500	11	18	2	2390	97	48	65	29 – 74
Northwestern	2984	167	405	132	7	2257	16	62	65	21 – 98
Vanderbilt	2432	24	153	23	3	2205	24	47	69	27 – 97
<i>SC:BCM-HGSC</i>	<i>13972</i>									
<b>TOTAL</b>	<b>21915</b>	<b>1666</b>	<b>3279</b>	<b>1497</b>	<b>77</b>	<b>14480</b>	<b>916</b>	<b>53.0</b>	<b>62</b>	<b>4 – 104</b>



**Table 2:**

Actionable incidental findings variants identified incidentally in 5 participants.

Gene	DNA change	Protein consequence	Interpre-tation	Participant count	rsID	ClinVar Variant ID
<i>BRCA2</i>	c.5946del	p.Ser1982Argfs*22	P	17	rs80359550	9325
<i>MYBPC3</i>	c.3628-41_3628-17del	splice	LP	13	rs36212066	177677
<i>APOB</i>	c.10580G>A	p.Arg3527Gln	P	12	rs5742904	17890
<i>BRCA1</i>	c.68_69del	p.Glu23Valfs*17	P	9	rs80357914	17662
<i>RET</i>	c.2410G>A	p.Val1804Met	P	7	rs79658334	37102
<i>KCNQ1</i>	c.1085A>G	p.Lys362Arg	LP	6	rs12720458	52953
<i>MYBPC3</i>	c.1504C>T	p.Arg502Trp	P	6	rs375882485	42540
<i>PMS2</i>	c.137G>T	p.Ser46Ile	P	5	rs121434629	9245
<i>RYR1</i>	c.1840C>T	p.Arg614Cys	LP/P <sup>‡</sup>	5	rs118192172	12964
<i>PKP2</i>	c.2146-1G-C	splice	P	5	rs193922674	6756

<sup>‡</sup>Sequencing centers differ in their interpretation of this variant

**Table 3:**  
**Incidental finding rates stratified by self-reported participant race/ethnicity.**

Self-reported race/ethnicity derives from multiple-choice selection of US Census-standardized terms and does not fully capture the complexity of participant racial identity.

Self-reported Race/ethnicity	Number of participants	Number of Incidental Findings	Fraction of LP findings (%)	Incidental Findings Rate % (95% CI)
American Indian, Alaska Native, or Pacific Islander	77	1	100	1.30 (0.23 – 6.97)
Hispanic or Latinx	1666	33	42	1.98 (1.41 – 2.77)
Black or African-American	3279	75	40	2.29 (1.83 – 2.86)
Asian	1497	41	59	2.74 (2.02 – 3.69)
White	14480	450 <sup>†</sup>	33	3.10 (2.84 – 3.40)
Unknown / Not reported	916	55	36	6.00 (4.64 – 7.73)

<sup>†</sup> 65 of these are *HFE* p.C282Y homozygotes (NM\_000410.3:c.845G>A)