HHS Public Access

Author manuscript

J Investig Dermatol Symp Proc. Author manuscript; available in PMC 2018 March 21.

Published in final edited form as:

J Investig Dermatol Symp Proc. 2018 January ; 19(1): S57–S61. doi:10.1016/j.jisp.2017.10.006.

Big Data Reveal Insights into Alopecia Areata Comorbidities

Chean Ping Lim[1], **Rachel K. Severin**[2], and **Lynn Petukhova**[1,2]

[1]Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, New York, USA

[2]Department of Dermatology, College of Physicians and Surgeons, Columbia University, New York, New York, USA

Abstract

Autoimmune diseases create a substantial burden of disease, and alopecia areata is among the more prevalent forms. Comorbidities are medical conditions that tend to occur together and may provide etiologic insights, suggest novel therapeutic strategies, and help patients and family members understand the risk of other health conditions. It is well established that having one autoimmune disease increases risk for others because of an underlying shared biology. Precision medicine initiatives are creating vast amounts of data that allow us to efficiently identify comorbidities. A survey across various datasets suggests that patients with autoimmune disease, and patients with alopecia areata in particular, may have comorbid neuropsychiatric and metabolic conditions.

Precision medicine, big data, and comorbidity

Precision medicine initiatives aim to improve efficiencies in our health care system by deriving new biological and clinical knowledge from big data. The initial impetus for precision medicine came from the discovery that tumors could be molecularly subtyped with genome-wide gene expression data, and that these data capture information about clinically relevant biological differences between the tumors (Perou et al., 2000). Genomic data not only provide greater diagnostic resolution than classification schemes based on clinical observations and histopathology, but also provide insight about disease mechanisms, prognosis, and drug response (Sorlie et al., 2001). Today, molecular typing of tumors is routinely conducted and helps guide the efficient development of treatment plans (e.g., BCR/ABL, EGFR, HER2, ALK, Ras, MET, Raf). Cancers that harbor drug-gable molecular alterations are targeted with precision and accuracy, and patients are spared unnecessary treatments.

Disease subtypes, created by clinically important biological differences among patients who share a common diagnosis, are not unique to cancer. Autoimmune diseases (AD) represent a large class of disorders characterized by an aberrant interaction between the immune system

Correspondence: Lynn Petukhova, College of Physicians & Surgeons, Columbia University, Russ Berrie Medical Science Pavilion, 1150 St. Nicholas Avenue, Room 303B, New York 10032, USA. lynn.petukhova@columbia.edu.

CONFLICT OF INTEREST

The authors state no conflict of interest.

and a targeted tissue that results in an expansion of immune cell populations that damage or destroy an end organ. Epidemiologic, genomic, and pharmacologic studies all indicate that there is a complex causal structure underlying AD, whereby each diagnosis comprises disease subtypes (Banchereau et al., 2016; Virgin and Todd, 2011), and some disease mechanisms are shared between AD with very different clinical presentations (Betz et al., 2015; Farh et al., 2015). Such mechanistic links provide drug repurposing opportunities, which are especially critical for AD given that so many of the disorders lack any evidence-based treatments. Our genetic research in alopecia areata (AA) provides an illustrative example, as mechanistic connections to rheumatoid arthritis first revealed through our genome-wide association study provided the initial rationale to repurpose drugs targeting the costimulatory and JAK-signal transducer and activator of transcription signaling pathways and led to clinical trials within only a few years (Mackay-Wiggan et al., 2016; Petukhova et al., 2010; Xing et al., 2014).

Disease mechanisms are at the crux of precision medicine, allowing us to resolve disease subtypes and identify drug repurposing opportunities (Figure 1). Thus, in laying out a roadmap for precision medicine, the National Institutes of Health has proposed the use of data science to extract high-resolution definitions of disease mechanisms from big data, advocating interrogation of diverse and complementary sources of data, including human genome data and electronic health records (EHR), among others (Collins, 2015). Although each dataset may be subject to different sources of bias that may warrant cautious interpretation of analytic results (Hripcsak and Albers, 2013; Ross et al., 2013), consistency of evidence across diverse datasets suggests validity. Large-scale data provide an opportunity to make discoveries that are unbiased by our previous assumptions about causes of disease.

Comorbidities can provide important etiologic insights, suggesting the presence of disease mechanisms that link clinical entities, and providing a classification scheme for subtyping patients. For example, if we consider Figure 1, disease mechanism 1 is a shared cause of diagnoses A and B. Thus, patients who harbor that mechanism are at an increased risk of both diseases, and conversely, we can increase our power to detect components of that disease mechanism by assembling a cohort of patients who are comorbid for those two diseases. Historically, studies of comorbidities have been limited to small cohorts testing specific hypotheses. For example, it has been known that having one AD increases the risk of other AD, and so the vast majority of comorbidity studies in AA have focused on other autoimmune or inflammatory conditions (e.g., Barahmani et al., 2009).

Big data insights into AA comorbidities

One of the first large-scale studies performed in EHR examined the codistribution of 161 International Classification of Disease (ICD) codes, including AA, among 1.5 million patient records (Rzhetsky et al., 2007). This study found enrichment of 34 ICD codes among patients with AA, including not only immune-related conditions, but also several neuropsychiatric disorders including epilepsy, depression, migraines, attention deficit, and bipolar disorders. Metabolic conditions included type 2 diabetes, and disorders of lipid metabolism and aromatic amino acid metabolism. To further investigate disease comorbidities in AA, we next leveraged several large publicly available data sources and

discovered consistent evidence for comorbidities not only involving immune system dysfunction (e.g., other autoimmune and inflammatory diseases and deficiencies in immune response to pathogens), but also neuropsychiatric and metabolic disorders (Table 1).

Phenome-wide association studies are conducted in cohorts of patients for whom EHR are linked to genetic data, testing for associations of ICD codes with disease risk alleles previously identified in the genome-wide association studies. This method provides an unbiased high-throughput method for identifying pleiotropic effects of single nucleotide polymorphisms. An online catalog of phenome-wide association study results (phewas.org) contains 29 single nucleotide polymorphisms from 12 of the 14 genomic regions associated with AA (Betz et al., 2015; Petukhova et al., 2010) and implicates more than 1,000 ICD codes (Table 1). Known AA comorbidities such as hypothyroidism, rheumatoid arthritis, celiac disease, type 1 diabetes, and psoriasis are implicated by multiple regions. Interestingly, several of the disorders identified in the Rzhetsky et al. study are also captured in these data. For example, four regions implicate major depressive disorder, and five regions implicate type 2 diabetes. Although two regions implicate hyperlipidemia, five additional regions also implicate atherosclerosis. As these are both risk factors for cardiovascular disease, it is noteworthy that 11 of the 12 regions contain single nucleotide polymorphisms that confer risk for AA and cardiac disease. Also of interest, seven loci implicate hypertension, which was not included in the Rzhetsky study.

Many high-throughput screening assays of traits in mouse models include the model that develops spontaneous AA, C3H/HeJ, and are cataloged online (phenome.jax.org). It is possible to query the database to identify traits for which a particular strain is an outlier, with values that are more than two standard deviations beyond the trait mean. Several traits that could be related to some of the human comorbidities are identified in this way, including autoimmune, inflammatory, metabolic, and neuropsychiatric conditions (Table 1). These data further suggest a biological basis to comorbidities in AA that do not appear to be directly related to an aberrant immune response or hair follicle integrity.

Our group next pursued investigation of lipid metabolism. This trait was implicated in all of the examined datasets, suggesting that patients with AA may be at an increased risk of heart disease, as has been established for rheumatoid arthritis and psoriasis (Crowson et al., 2013; Gerdes et al., 2014). Using Columbia University Medical Center's EHR, we obtained clinical reports for lipid tests conducted for patients with an ICD code for AA, as well as several common dermatological conditions for comparison. We identified 752 patients with AA and 21,538 controls seen at Columbia University Medical Center between 2009 and 2014. Among the 22,290 subjects, lipid profiles from 145 patients with AA and 6,690 controls were available for analysis. We estimated the adjusted change in lipid levels in patients with AA using linear regression for four lipid profiles: total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein, and triglycerides. We found a significant increase of LDL in patients with AA compared with controls, but only for females (P = 0.01). On average, female patients who had a recorded diagnosis of AA have LDL levels that are increased by 8.16 mg/dl compared with controls, adjusting for age (β = 8.16, 95% CI: 1.84–14.48). However, we did not observe similar elevation in males. Data for LDL in females across all examined conditions are shown in Figure 2.

Although lipid dysregulation has not yet been extensively studied in the context of AA, there are several lines of evidence to suggest that elevated LDL is a plausible and clinically relevant comorbidity. Lipids and metabolic intermediates in the cholesterol biosynthesis pathway have been directly or indirectly linked to both immune responses and hair disorders. For example, cholesterol biosynthesis provides energy and raw materials required for the proliferation of cells of the immune system in response to activation (Getz and Reardon, 2014), and metabolic intermediates in the cholesterol biosynthetic pathway directly influence signaling pathways relevant to immune responses and hair growth, including JAK-signal transducer and activator of transcription and peroxisome proliferator-activated receptor signaling (Davies et al., 2016). Studies in mouse models and patients demonstrate that perturbations in the cholesterol biosynthesis pathway underlie some hair disorders (Stenn and Karnik, 2010; Evers et al., 2010). Finally, pharmacologic evidence and clinical trials in other autoimmune diseases, including rheumatoid arthritis, demonstrate that LDL-lowering therapies can mitigate risk and attenuate symptoms of autoimmune disease (Davies et al., 2016; Greenwood et al., 2006).

In AA, two small, uncontrolled pilot trials have investigated the use of simvastatin/ezetimibe to treat AA and have found conflicting results (Lattouf et al., 2015; Loi et al., 2016). Although it is not clear if the inconsistency could be due to differences in study design or limitations in power, it is intriguing to note that within the study that found a beneficial effect (Lattouf et al., 2015), the mean regrowth for women (71%) was greater than the mean regrowth for men (57%), supporting our identification of a gender effect in LDL levels among patients with AA (Figure 3). Although evidence for the efficacy of statins in AA is tenuous, research in other AD, along with a strong body of mechanistic data, warrants pursuit of future clinical trials in patients with AA, and the initiation of studies in the mouse model of AA (Davies et al., 2016; Greenwood et al., 2006).

Conclusions

Big data offer us an opportunity to perform research that is unbiased by previous assumptions about causal structure. Genome-wide association studies have shown us that such approaches can provide new insight with clinical relevance and can profoundly impact patient care. The promise for precision medicine is that the use of big data drawn from different sources can provide an increased resolution of disease mechanisms, enhancing our ability to identify clinically meaningful patient subtypes and mechanistic links between different diagnoses that motivate drug repurposing.

By interrogating several large datasets for comorbid conditions in AA, we have revealed preliminary evidence that suggests the complex causal structure underlying AA could also influence metabolic and neuropsychiatric traits and conditions. Each of these datasets is subject to various sources of bias, and they do not provide insight about the underlying causal order of observed co-occurrences. Nonetheless, consistency across the datasets suggests that further investigations are warranted. Furthermore, our observations support previous epidemiologic studies that have established links between psoriasis and obesity, lupus and neuropsychiatric disorders, and rheumatoid arthritis and heart disease and suggest

that future investigations should expand the scope of inquiry to include multiple autoimmune and inflammatory conditions, which our group has initiated.

Acknowledgments

Funding for the Summit and the publication of this article was provided by the National Alopecia Areata Foundation. Funding for this Summit was also made possible (in part) by a grant (1 R13AR071266) from the National Institute of Arthritis and Musculoskeletal and Skin Diseases (NIAMS). The views expressed in written conference materials or publications and by speakers and moderators do not necessarily reflect the official policies of the Department of Health and Human Services; nor does mention of trade names, commercial practices, or organizations imply endorsement by the U.S. Government.

This work was generously supported in part by funding from the National Alopecia Areata Foundation (to LP). We thank Dr Prakash Gorroochurn for biostatistical advice. We are grateful for support from 5U01HG008680 Columbia GENIE (Genomic Integration with EHR), P30AR069632 Columbia University Skin Disease Resource-Based Center (epiCURE), and P50AR070588 Alopecia Areata Center for Research Translation (AACORT).

References

Banchereau R, Hong S, Cantarel B, Baldwin N, Baisch J, Edens M, et al. Personalized immunomonitoring uncovers molecular networks that stratify lupus patients. Cell. 2016; 165:551–65. [PubMed: 27040498]

Barahmani N, Schabath MB, Duvic M, National Alopecia Areata R. History of atopy or autoimmunity increases risk of alopecia areata. J Am Acad Dermatol. 2009; 61:581–91. [PubMed: 19608295]

Betz RC, Petukhova L, Ripke S, Huang H, Menelaou A, Redler S, et al. Genome-wide meta-analysis in alopecia areata resolves HLA associations and reveals two new susceptibility loci. Nat Commun. 2015; 6:5966. [PubMed: 25608926]

Collins, FS. [accessed October 2015] ASHG 2015 Presidential Symposium. 2015. http://www.ashg.org/2015meeting/

Crowson CS, Liao KP, Davis JM III, Solomon DH, Matteson EL, Knutson KL, et al. Rheumatoid arthritis and cardiovascular disease. Am Heart J. 2013; 166:622–628.e1. [PubMed: 24093840]

Davies JT, Delfino SF, Feinberg CE, Johnson MF, Nappi VL, Olinger JT, et al. Current and emerging uses of statins in clinical therapeutics: a review. Lipid Insights. 2016; 9:13–29. [PubMed: 27867302]

Evers BM, Farooqi MS, Shelton JM, Richardson JA, Goldstein JL, Brown MS, et al. Hair growth defects in Insig-deficient mice caused by cholesterol precursor accumulation and reversed by simvastatin. J Invest Dermatol. 2010; 130:1237–48. [PubMed: 20090767]

Farh KK, Marson A, Zhu J, Kleinewietfeld M, Housley WJ, Beik S, et al. Genetic and epigenetic fine mapping of causal autoimmune disease variants. Nature. 2015; 518:337–43. [PubMed: 25363779]

Gerdes S, Osadtschy S, Buhles N, Baurecht H, Mrowietz U. Cardiovascular biomarkers in patients with psoriasis. Exp Dermatol. 2014; 23:322–5. [PubMed: 24660963]

Getz GS, Reardon CA. The mutual interplay of lipid metabolism and the cells of the immune system in relation to atherosclerosis. Clin Lipidol. 2014; 9:657–71. [PubMed: 25705263]

Greenwood J, Steinman L, Zamvil SS. Statin therapy and autoimmune disease: from protein prenylation to immunomodulation. Nat Rev Immunol. 2006; 6:358–70. [PubMed: 16639429]

Hripcsak G, Albers DJ. Next-generation phenotyping of electronic health records. J Am Med Inform Assoc. 2013; 20:117–21. [PubMed: 22955496]

Lattouf C, Jimenez JJ, Tosti A, Miteva M, Wikramanayake TC, Kittles C, et al. Treatment of alopecia areata with simvastatin/ezetimibe. Am Acad Dermatol. 2015; 72:359–61.

Loi C, Starace M, Piraccini BM. Alopecia areata (AA) and treatment with simvastatin/ezetimibe: experience of 20 patients. J Am Acad Dermatol. 2016; 74:e99–100. [PubMed: 27085249]

Mackay-Wiggan J, Jabbari A, Nguyen N, Cerise JE, Clark C, Ulerio G, et al. Oral ruxolitinib induces hair regrowth in patients with moderate-to-severe alopecia areata. JCI Insight. 2016; 1:e89790. [PubMed: 27699253]

Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, et al. Molecular portraits of human breast tumours. Nature. 2000; 406:747–52. [PubMed: 10963602]

Petukhova L, Duvic M, Hordinsky M, Norris D, Price V, Shimomura Y, et al. Genome-wide association study in alopecia areata implicates both innate and adaptive immunity. Nature. 2010; 466:113–7. [PubMed: 20596022]

Ross MG, Russ C, Costello M, Hollinger A, Lennon NJ, Hegarty R, et al. Characterizing and measuring bias in sequence data. Genome Biol. 2013; 14:R51. [PubMed: 23718773]

Rzhetsky A, Wajngurt D, Park N, Zheng T. Probing genetic overlap among complex human phenotypes. Proc Natl Acad Sci USA. 2007; 104:11694–9. [PubMed: 17609372]

Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc Natl Acad Sci USA. 2001; 98:10869–74. [PubMed: 11553815]

Stenn KS, Karnik P. Lipids to the top of hair biology. J Invest Dermatol. 2010; 130:1205–7. [PubMed: 20393475]

Virgin HW, Todd JA. Metagenomics and personalized medicine. Cell. 2011; 147:44–56. [PubMed: 21962506]

Xing L, Dai Z, Jabbari A, Cerise JE, Higgins CA, Gong W, et al. Alopecia areata is driven by cytotoxic T lymphocytes and is reversed by JAK inhibition. Nat Med. 2014; 20:1043–9. [PubMed: 25129481]

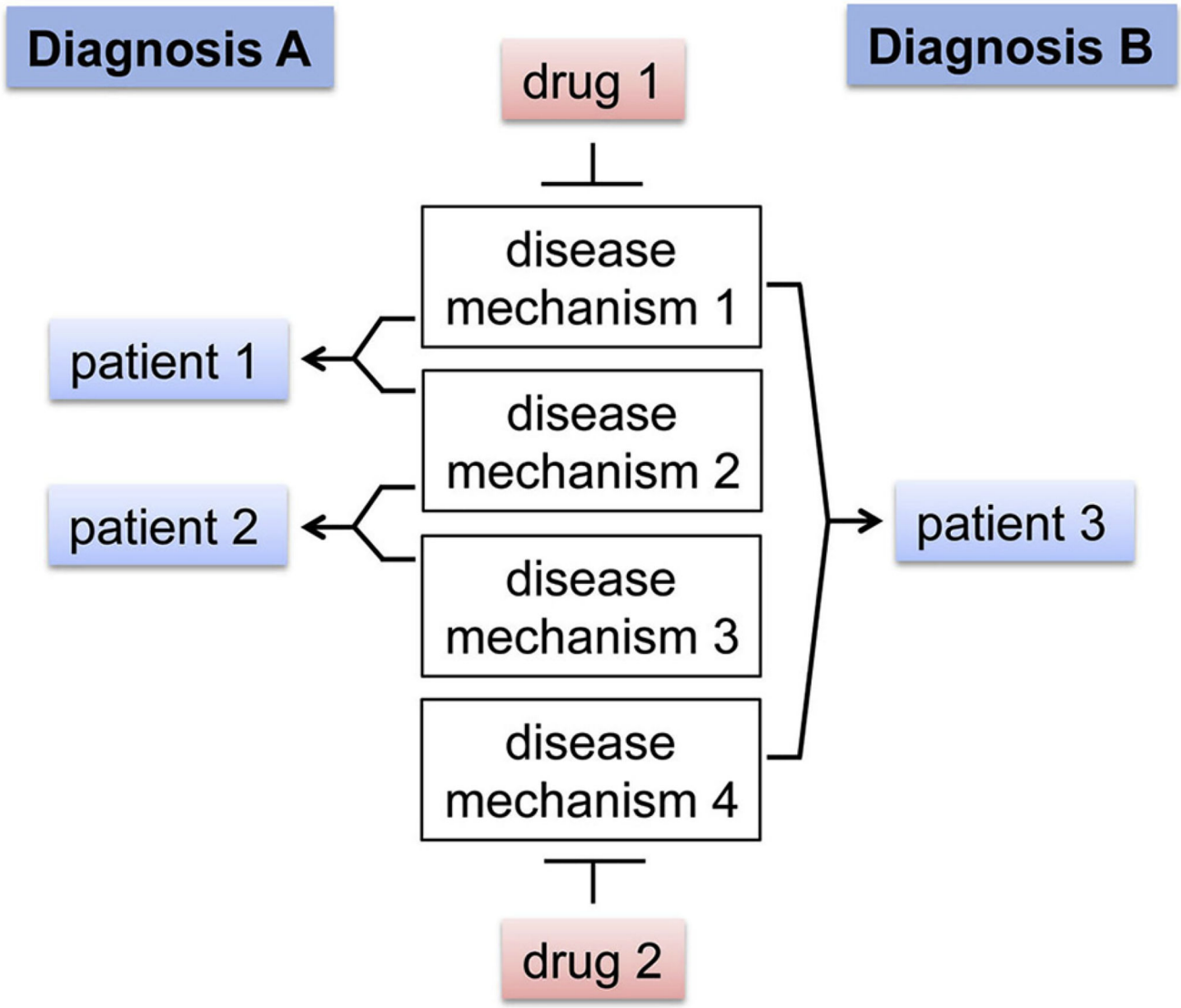

Figure 1.

Causal structure of two complex diseases that have a clinically relevant mechanistic link. Different configurations of disease mechanisms can give rise to subtypes of disease (patients 1 and 2 share diagnosis A), or to diseases that have different clinical presentations, warranting unique diagnoses (patient 3 has diagnosis B). This causal structure has important therapeutic implications. For example, drug 1 would fail to illicit a therapeutic response for patient 2, but could be repurposed to treat patients with diagnosis B (patient 3). Thus, disease mechanism 1 provides a clinically relevant causal link between diagnoses A and B. Observed variables are in colored boxes.

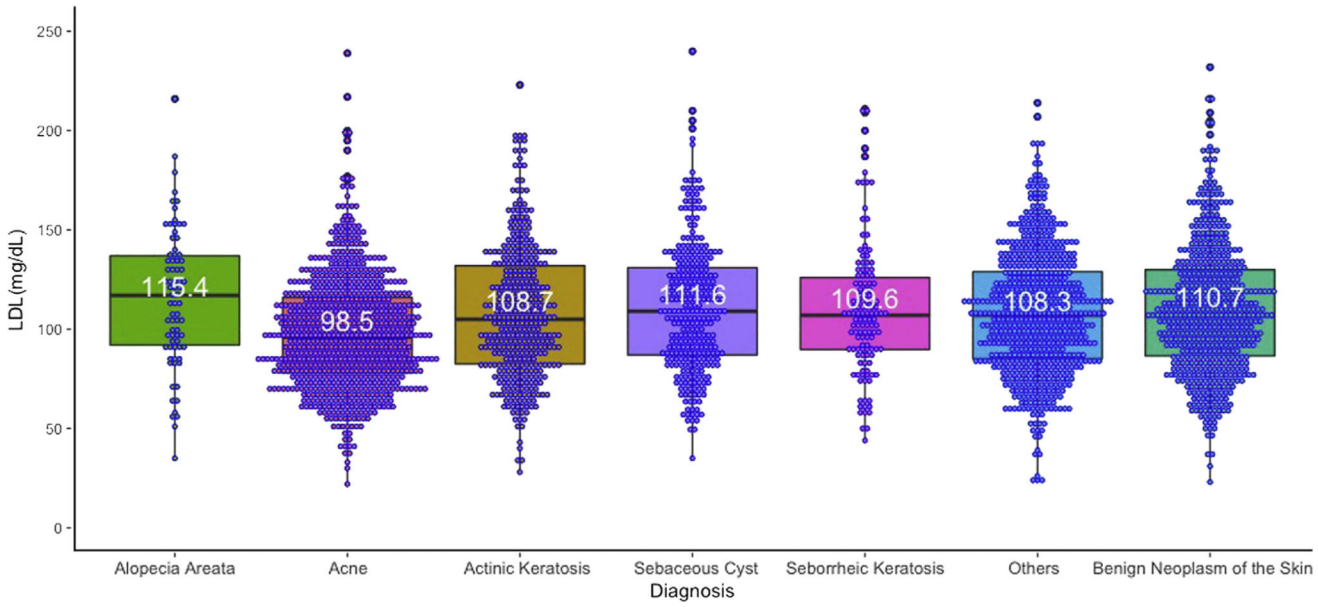

Figure 2.

LDL levels in female patients. Patients are grouped by dermatological diagnoses, excluding patients with multiple dermatological diagnoses. Box plots are generated for each group, with lower and upper border representing the 25th and 75th percentile, respectively; the middle line within each box plot represents the median level of LDL (mg/dl). Dot plots are superimposed to show individual measurements. Mean levels of LDL (mg/dl) are displayed over the dot and box plots. Sample size for each group is shown in legends. LDL, low-density lipoprotein.

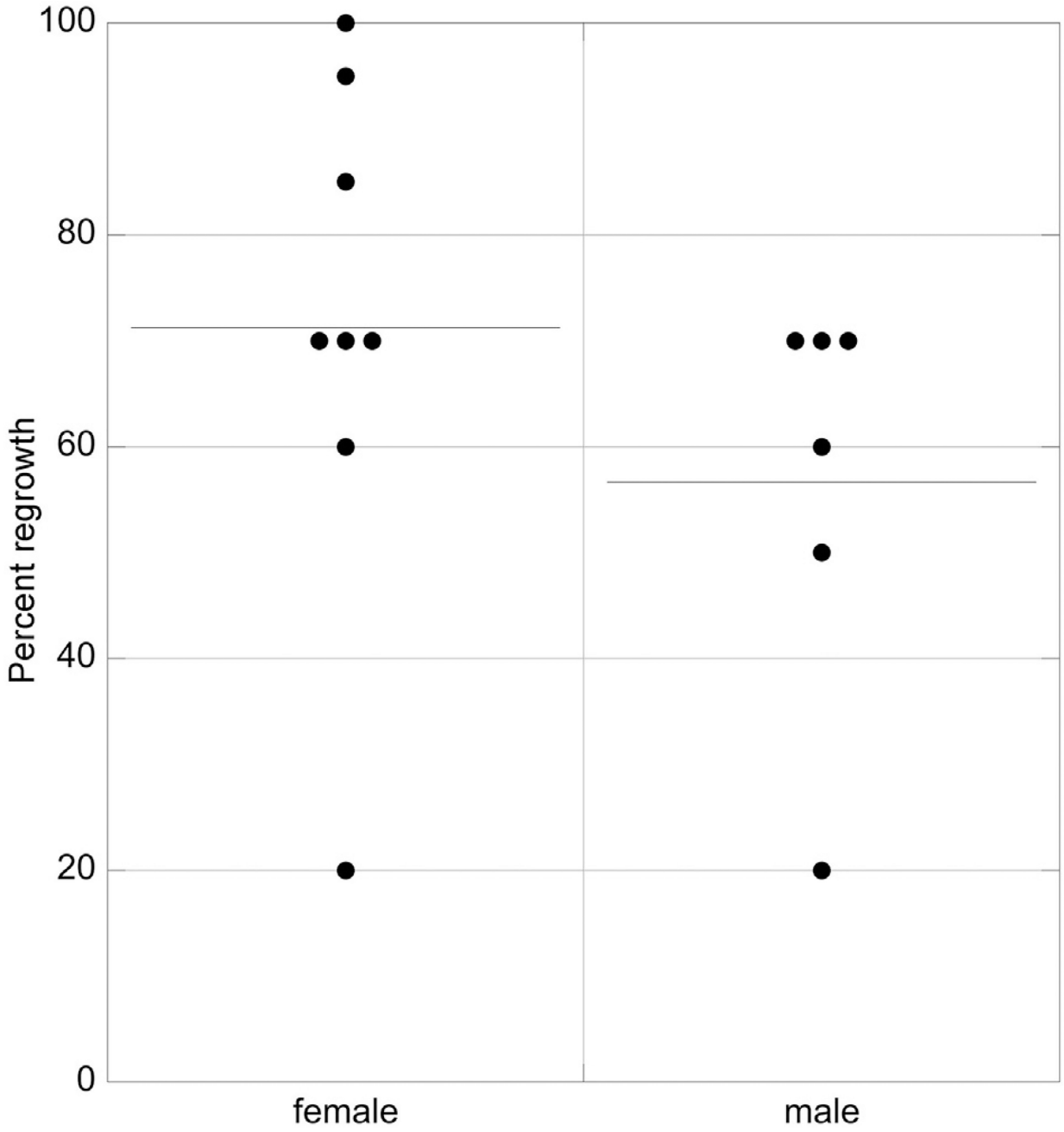

Figure 3.

Percent hair regrowth reported in a clinical trial of simvastatin/ezetimibe reported in Lattouf et al. (2015). Data reported in Table 1 of Lattouf et al. are plotted by gender. Although the difference between genders is not statistically significant (Mann-Whitney $P = 0.18$), the sample size (N = 14) has very limited power to detect differences. The direction of effect is consistent with our observation of higher LDL levels in female patients with AA and no difference in male patients relative to controls. Although this trial did not report LDL levels,

our data suggest that future trials should record and monitor cholesterol levels. LDL, low-density lipoprotein.

Table 1

Alopecia areata comorbidities suggested by big data

Disease/trait category	ICD co-occurrence (Rzhetsky et al. 2007)	PheWAS (phewascatalog.org)	C3H/HeJ (phenome.jax.org)
Autoimmune	Rheumatoid arthritis	Rheumatoid arthritis	Prone to colitis
	Multiple sclerosis	Multiple sclerosis	Thyroid hormone abnormalities
	Systemic lupus Erythematosus	Systemic lupus Erythematosus	
	Type 1 diabetes mellitus	Type 1 diabetes mellitus	
	Psoriasis	Psoriasis	
		Ulcerative colitis	
		Grave's disease, Thyroiditis	
Inflammatory Metabolic	Allergic rhinitis	Allergic rhinitis	Elevated total cholesterol
	Hypersensitivity angiitis	Asthma	Elevated phospholipids
	Disorders of lipid metabolism	Disorders of lipid metabolism	Decreased circulating alanine transaminase
	Type 2 diabetes mellitus	Type 2 diabetes mellitus	
	Cholelithiasis		Elevated heme oxygenase
	AA metabolism (aromatic) acanthosis nigricans		
Neuropsychiatric	Attention deficit disorder	Attention deficit disorder	Prone to impulsivity, anxiety
	Epilepsy	Epilepsy	Absence seizures (*Gria4*^{spkw})
	Depression		Attenuated response to tactile and thermal stimulation
	Bipolar disorder		
	Migraine		Disruptions in social behavior
Cardiovascular		Atherosclerosis	Abnormal ECG findings
		Myocardial infarction	
		Blood pressure	
		Heart failure	

Three datasets were queried to identify conditions and traits that could be associated with AA. The table reports only conditions related to the four categories of focus in our commentary. Additional traits may be found in each source. ICD co-occurences were identified in an analysis of 161 ICD codes in 1.5 million patient records (Rzhetsky et al, 2007). We identified 28 SNPs associated with AA in our recent meta-analysis in the PheWAS catalogue and extracted all reported ICD code associations. For the C3H/HeJ, we manually curated a list of all traits for which this strain was reported to be an outlier.