Endotyping in Heart Failure
— Identifying Mechanistically Meaningful Subtypes of Disease —

Lusha W. Liang, MD; Yuichi J. Shimada, MD

Endotyping is an emerging concept in which diseases are classified into distinct subtypes based on underlying molecular mechanisms. Heart failure (HF) is a complex clinical syndrome that encompasses multiple endotypes with differential risks of adverse events, and varying responses to treatment. Identifying these distinct endotypes requires molecular-level investigation involving multi-“omics” approaches, including genomics, transcriptomics, proteomics, and metabolomics. The derivation of these HF endotypes has important implications in promoting individualized treatment and facilitating more targeted selection of patients for clinical trials, as well as in potentially revealing new pathways of disease that may serve as therapeutic targets. One challenge in the integrated analysis of high-throughput omics and detailed clinical data is that it requires the ability to handle “big data”, a task for which machine learning is well suited. In particular, unsupervised machine learning has the ability to uncover novel endotypes of disease in an unbiased approach. In this review, we will discuss recent efforts to identify HF endotypes and cover approaches involving proteomics, transcriptomics, and genomics, with a focus on machine-learning methods.

Key Words: Artificial intelligence; Endotypes; Genomics; Heart failure; Proteomics; Transcriptomics

Heart failure (HF) affects over 60 million people worldwide and is a leading cause of morbidity and mortality.1 The diagnosis of HF is based on an assessment of clinical features and has historically been divided by left ventricular (LV) ejection fraction (EF) into 2 categories: HF with reduced EF (HFrEF) and HF with preserved EF (HFpEF).2 However, this simplistic approach fails to capture the underlying heterogeneity of the etiology and pathophysiology of HF, and more nuanced approaches to subtyping HF are needed.3 The limitation of this approach is perhaps best exemplified by the lack of effective treatments for patients with HFpEF. Conversely, HFrEF now exists in an era of multiple therapies, and successful strategies for guiding individualized treatment in HFrEF will also need to rely on a deeper understanding of the pathophysiology of HF subtypes beyond classification by EF alone.4 Finally, there has been growing interest in the entity of HF with mid-range EF (HFmrEF), which also encompasses a wide array of patients, from those with HF with recovered EF to those who progress to HFpEF or HFrEF.2 5

The most common forms of HF result from a complex interplay between multiple genetic and environmental factors that lead to myocardial dysfunction, and thus prognostication based on clinical data alone can be challenging.6 Numerous prognostic markers of death and/or hospitalization have been identified in patients with HF; however, their clinical applicability has remained limited.2 This highlights the need for new methodologies in the classification and characterization of HF.

Such strategies include the use of large-scale studies of molecular-level abnormalities in HF encompassing a multi-“omics” approach (e.g., genomics, proteomics, transcriptomics, and metabolomics). One challenge in the integrated analysis of high-throughput “omics” data is that it requires the ability to handle “big data”, a task for which machine learning is well suited.7 In this review, we will discuss recent efforts to identify subtypes of disease in HF and cover approaches involving proteomics, transcriptomics, and genomics, with a focus on ML methods.

Supervised ML and Unsupervised Clustering

ML is a branch of artificial intelligence (AI) consisting of pattern-recognition algorithms used to define relationships between objects.8 ML is typically subdivided into supervised and unsupervised learning (Figure 1A).8 Supervised ML focuses on making predictions after training on a selected dataset.9 Unsupervised ML was designed to determine the intrinsic structure of a dataset.8 Although supervised ML is useful, both in its own right and as a method for validating the results of unsupervised ML, the use of supervised ML models relies on pre-existing classifications of HF and therefore is not suited to the discovery of novel subtypes of disease.

Cluster analysis is an important component of unsupervised ML whose primary goal is to find related groups in a...
Figure 1. Commonly used machine learning (ML) methods in medicine. (A) Supervised vs. unsupervised ML. Supervised ML is used for classification and regression problems whereas unsupervised ML is used to determine the intrinsic structure of a dataset and uncover distinct clusters of data. Commonly used supervised ML algorithms include decision tree-based methods such as random forest algorithms, neural networks, and support vector machines. Clustering techniques include k-means clustering and partitioning around medoids, hierarchical clustering, and model-based clustering. (B) Datasets that are well suited vs. not well suited for analysis using k-means clustering analysis. (Left) A dataset that forms circular clusters is well suited for analysis using k-means clustering. (Right) A dataset that forms noncircular clusters may not be well suited for analysis using k-means clustering. (C) Dendrogram formed by use of a hierarchical clustering algorithm. (D) Use of gap statistics to determine optimal number of clusters in unsupervised ML. Dividing this example dataset into 4 clusters maximizes the gap statistic.
Endotyping in HF

nogroups” in HFpEF, Shah et al relied on deep clinical phenotyping and used hierarchical clustering techniques to identify 3 clusters of patients: a younger group with lower B-type natriuretic peptide (BNP) levels, an older group with chronic kidney disease (CKD), and a group characterized by multiple comorbidities, including obesity and diabetes. In a validation cohort, phenogroup membership was independently associated with adverse outcomes, with patients in the phenogroup characterized by CKD having the highest risk of adverse outcomes. A number of subsequent studies have further validated this approach in patients with HFpEF.

In these studies, the phenogroup of patients characterized by a heavier burden of comorbidities, including CKD, was again found to have the highest risk of adverse outcomes.

Focusing instead on HFrEF, Ahmad et al used unsupervised ML clustering techniques to identify 4 distinct clusters of patients in a group of 1,619 participants. Although LVEF was not significantly different across all 4 patient clusters, there were significant differences between the clusters in clinical characteristics, outcomes, and responses to therapy. In another study examining patients with HFrEF undergoing cardiac resynchronization therapy (CRT), unsupervised ML detected 2 phenogroups of patients who were found to respond substantially better to CRT therapy than patients in other phenogroups.

Applying TDA to echocardiographic features of LV structure and function, Tokodi et al generated a network with 4 distinct regions, with patients in each region demonstrating significant differences in major adverse cardiac event (MACE)-related rehospitalization and death. Collectively, these studies using clinically available data demonstrate the phenotypic diversity of HF and serve as a proof-of-concept that unsupervised ML can provide clinically meaningful classifications of HF that can aid in optimizing management and treatment strategies. However, derivation of HF subtypes using only clinical data does not reflect the underlying dataset without the use of a response variable. Multiple clustering algorithms exist, each with its own set of benefits and drawbacks. One of the simplest clustering algorithms is k-means clustering wherein data points are organized around centroids, and the objective is to identify the configuration that minimizes the distance, or dissimilarity, of the data points to their closest selected centroid. The k-means clustering functions best when the dataset forms circular clusters. Hierarchical agglomerative clustering is another commonly used algorithm, in which each data point begins as a singleton cluster and pairs of clusters are then successively merged until all data points are eventually contained in 1 large cluster. The result is a tree-based representation of the data points, called a dendrogram. However, there is no systematic guidance as to where to cut the dendrogram to form clusters.

Indeed, a major challenge in cluster analysis is the estimation of the optimal number of clusters. Gap statistics is a commonly used method to estimate the number of clusters in a dataset by comparing the change in within-cluster dispersion with that expected under the null distribution. In addition, a model-based clustering algorithm can be used in which each cluster is mathematically represented by a parametric distribution (e.g., Gaussian) so that the entire dataset is modeled by a mixture of these distributions; cluster assignment is then made based on maximization of a penalized likelihood. Finally, topological data analysis (TDA) refers to a collection of innovative statistical methods that provide insight into the geometric structure of data and generate useful visualizations of complex datasets. TDA can be used in an unsupervised ML pipeline to cluster similar patients into nodes.

Approach to Heterogeneity Using Clinical Data: “Phenogrouping”

In the first study to use unsupervised ML to classify “phe-
Endotypes can identify groups of patients who may respond differently to treatment based on shared dysfunctional molecular pathways. This is particularly important in HF because it encompasses multiple endotypes that may have differential risks of adverse events and varying responses to treatment. The previous studies using ML to identify phenogroups have demonstrated this concept using clinical data. However, identifying endotypes takes this concept one step further into molecular-level investigations. Studies using omics data and unsupervised clustering approaches can provide unbiased molecular evidence for derivation of endotypes. These studies also have the potential to reveal novel insights into disease pathophysiology as well as facilitating the development of individualized treatment.

The Emerging Concept of “Endotyping”

HF is a complex syndrome, and diagnosis is based on a comprehensive clinical assessment. Severity of symptoms can be further classified using the New York Heart Association functional classes and the American College of Cardiology/American Heart Association staging system. However, these classifications reflect downstream consequences of myocardial dysfunction rather than underlying molecular and cellular mechanisms of disease. “Endotyping” is an emerging concept in which diseases are classified into distinct subtypes based on underlying molecular mechanisms (Figure 2). Endotypes can identify groups of patients who may respond differently to treatment based on shared dysfunctional molecular pathways. This is particularly important in HF because it encompasses multiple endotypes that may have differential risks of adverse events and varying responses to treatment. The previous studies using ML to identify phenogroups have demonstrated this concept using clinical data. However, identifying endotypes takes this concept one step further into molecular-level investigations. Studies using omics data and unsupervised clustering approaches can provide unbiased molecular evidence for derivation of endotypes. These studies also have the potential to reveal novel insights into disease pathophysiology as well as facilitating the development of individualized treatment.
a statement addressing the need for more targeted approaches to therapy for patients with HFpEF, and recommended that “successful strategies for how to guide medical therapy for patients will have to rely on the emergence of health data science and a deeper understanding of fundamental biology, pathophysiology, genomics, and phenotyping.” In the next few sections, we will discuss recent efforts to identify mechanistically meaningful endotypes of HF and identify molecular markers of different disease states using proteomics, transcriptomics, and genomics.

Endotyping Using Proteomics Profiling

Proteomics profiling involves the measurement of hundreds to thousands of proteins simultaneously. Many proteins that are involved in signaling pathways in the heart can be detected in peripheral plasma through the use of proteomics profiling. Applying unsupervised ML to proteomics profiling data can lead to the identification of clusters of patients with unique proteomics profiling signatures, thus illuminating underlying mechanisms of disease and allowing for the derivation of molecularly distinct endotypes (Table).

Woolley et al examined a panel of 363 proteomic biomarkers from 429 patients with HFpEF and used unsupervised ML to identify 4 distinct endotypes with the following clinical characteristics: a younger group with lower N-terminal-proB-type natriuretic peptide (NT-proBNP) levels, an older group with CKD, a group with multiple comorbidities, and a group with significant coronary artery disease (CAD). Interestingly, the clinical characteristics of the endotypes were very similar in this study using proteomics data compared with the Shah study using clinical phenotyping. The group with the highest prevalence of CKD was associated with worse outcomes in both studies. Pathway analysis revealed upregulation of inflammatory pathways in the endotype characterized by CKD, as well as upregulation of pathways implicated in cell proliferation regulation and cell survival in the endotype characterized by ischemia.

Stienen et al performed a similar study, applying unsupervised ML techniques to 392 patients with HFpEF using a panel of 415 proteomic biomarkers. Their analysis identified 2 distinct endotypes, with patients in 1 endotype experiencing higher rates of cardiovascular death and hospitalization. Pathway analysis revealed upregulation of pathways involved in immune system activation, signal transduction cascades, cell interactions, and metabolism in the endotype with worse outcomes. Taken together, both studies demonstrate the heterogeneity of HFpEF and highlight potential future targets for investigation and development of mechanistically directed therapies. In addition, the proteomic biomarkers and pathways identified in both studies may help guide selection of patients for clinical trials in HFpEF who are more likely to benefit from a particular therapy based on their underlying pathophysiology.

Endotyping via ML-based analysis of proteomic data has been applied to HFrEF as well. Tromp et al applied unsupervised ML to a panel of 92 proteomic biomarkers in patients with HFrEF and identified 6 distinct endotypes with marked differences in clinical characteristics, outcomes, and response to medical therapy. Notably, 1 particular endotype did not derive benefit from β-blocker treatment despite being indistinguishable from the other patients with HFrEF based on clinical characteristics alone, demonstrating the added value of biomarker analysis and endotyping. A limited number of proteomic biomarkers could adequately discriminate patient endotype membership in this study, suggesting that such a panel of biomarkers could be used with relative ease to determine endotypes in a clinical setting.

One major challenge in the interpretation of proteomics profiling is the difficulty in assessing whether the differentially regulated pathways between endotypes are causal or secondary to disease progression. Integration with data from genomics and transcriptomics could further clarify the significance of pathways identified in proteomics.

Endotyping Using Transcriptomics

Transcriptomics involves the study of ribonucleic acid (RNA) transcripts that are produced by the genome using high-throughput methods such as microarray analysis. Comparison of transcriptomes allows the identification of genes that are differentially expressed in different cell populations, disease states or in response to different treatments. Transcriptomics has the potential to refine diagnostic and prognostic accuracy in a number of diseases, and has found considerable success in oncology. These techniques have not been developed as robustly in HF, although several studies have shown promise in correlating RNA transcript levels with different HF disease subtypes.

Kittleson et al used gene expression microarrays of myocardial samples obtained from patients with endstage HF at the time of transplantation or LV assist device implantation to develop a 90-gene panel on a training dataset; this prediction panel was then applied to a separate group of patient samples (test set). Using supervised principal components clustering techniques, the prediction panel was able to distinguish between patients with ischemic cardiomyopathy (ICM) and non-ischemic cardiomyopathy (NICM) with 100% sensitivity and specificity. Interestingly, when applied to patients with newly diagnosed cardiomyopathy, the prediction panel performed perfectly in NICM but only identified 1 of 3 ICM samples correctly. This suggests that patients with ICM experience greater changes in gene expression as the disease progresses when compared with patients with NICM, and emphasizes the need for stage-specific prediction profiles.

RNA-sequencing (RNA-seq) is a newer approach for transcriptome profiling and allows for an unbiased survey of the entire transcriptome. RNA-Seq has a greater dynamic range than microarrays, which can be susceptible to nonspecific hybridization and saturation biases. One study evaluated RNA-seq data from myocardial samples of 6 patients, comprising 1 patient with ICM, 2 patients with dilated cardiomyopathy (DCM), and 3 patients with non-failing (NF) hearts. Genes that were globally differentially expressed were then used as feature vectors to classify 313 individuals with microarray data using a k-means clustering algorithm. Remarkably, based on the feature vectors of only 6 patients, Liu et al demonstrated high accuracy in classifying patients between ICM and NF, as well as between DCM and NF. This study identified genes with distinct expression patterns between failing and NF hearts and found that these detailed expression profiles had the ability to distinguish different disease states.

Beyond gene expression profiles of messenger RNA (mRNA), many noncoding RNA molecules (e.g., microRNAs (miRNAs), long noncoding RNAs (lncRNAs), and circular RNAs) have been found to play an important role in
the regulation of gene transcription, epigenetics, and posttranscriptional mRNA processing and can be detected noninvasively in peripheral plasma (Figure 3). Circulating concentrations of noncoding RNA molecules vary in response to an array of acute and chronic disease states. miRNAs in particular are attractive candidates for biomarkers, given their stability in stored samples. Several studies have found that miRNA signatures can differentiate between patients with HFpEF vs. HFrEF, as well as between patients with dyspnea from HF vs. those with dyspnea from chronic obstructive pulmonary disease. In combination with NT-proBNP levels, the addition of miRNA panels further improves the accuracy of classification of patients with and without HF. In a study of 2,203 patients with chronic HF, higher levels of miR-1254 and miR-1306-5p were associated with higher risk of all-cause death and HF hospitalization, although hazard ratios were modest. Importantly, most of these studies utilized small panels of miRNA (4–12) to achieve high discriminative accuracy, suggesting that they could be attractive candidates for biomarkers in the clinical setting.

The lncRNAs, which are often defined as noncoding protein transcripts larger than 200 nucleotides, are also potential markers of cardiac dysfunction. They have been found to be independent predictors of diastolic function and remodeling in patients with diabetic cardiomyopathy, and have also been found to have value in predicting response to therapy in a trial of pioglitazone in patients with HF and diabetes. This has important implications in the use of transcriptomics as a means to guide treatment. Future efforts in transcriptomics will need to focus on improving standardization of microarray platforms and RNA-seq methods, as well as statistical data handling, to ensure that results are valid and generalizable.

Endotyping Using Genomics

Familial monogenetic cardiomyopathies are generally organized into several major phenotypic categories: hypertrophic, dilated, arrhythmogenic, restrictive, and LV non-compaction cardiomyopathy. However, even within these categories there is significant heterogeneity in the underlying etiology and clinical manifestations of disease. Moreover, many patients with cardiomyopathies have negative genetic testing, and even in those with known pathogenic variants, penetrance is often incomplete. In general, genetic testing for patients with cardiomyopathies is recommended when there is a family history of cardiomyopathy and cascade screening of at-risk family members is feasible and desired. The majority of circumstances, identification of a pathogenic variant does not alter treatment or risk stratification, because there is a lack of robust genotype–phenotype associations. A notable exception is patients with arrhythmogenic cardiomyopathy and a mutation in LMNA, FLNC or PLN. Mutations in these genes are associated with a higher risk of life-threatening arrhythmias, so for these patients there is a Class IIa recommendation for primary prevention with an implantable cardioverter defibrillator (ICD). In addition, patients with DCM and LMNA or SCN5A mutations may similarly be considered for primary prevention ICD.

Apart from the familial monogenic cardiomyopathies, epidemiologic studies have shown that genetic predisposition does still play a role in all-cause HF. A study evaluating the HF status of adoptees compared with their adoptive and biologic parents estimated the heritability of HF to be 26%. Genome-wide association studies (GWAS) have aimed to find associations between common single-nucleotide polymorphisms (SNPs) and HF using SNP arrays. In a GWAS of 47,309 patients with all-cause HF and 930,014 controls, Shah et al identified 12 independent variants at 11 genomic loci that were associated with HF. Despite the size of the study, only a modest number of genetic associations were identified and cumulative heritability was estimated at 9%, suggesting that an important component of HF heritability may be more attributable to specific disease subtypes than components of a final common pathway.

Indeed, in a prior study, GWAS was initially performed...
Endotyping in HF

comere protein titin, presents a unique challenge because it undergoes extensive alternative splicing to produce multiple isoforms. 

TTN mutations can cause DCM, and heterozygous mutations that truncate full-length titin (titin-truncating variants: TTNtv) are the most common genetic cause of familial DCM. In addition, TTNtv also occurs in about 2% of individuals without overt cardiomyopathy. In a study combining genetic, transcriptomic, proteomic, and clinical data, Roberts et al found that the clinical significance of TTNtv is largely determined by exon usage and variant location. In particular, TTNtv in exons with proportion-spliced-in (PSI) greater than 0.9 are much more likely to be pathogenic. Patients with DCM and high-PSI TTNtv were found to have earlier onset of HF, arrhythmias, and death than other patients with DCM. Thus, patients with TTNtv DCM may represent a higher-risk subtype of DCM who may benefit from a lower threshold for ICD therapy.

With the expansion of biobanks linking large-scale genomic sequence data to electronic health records (EHR), genomics-first studies have emerged. These studies first identify patients with a variant of interest, then use EHR data to associate clinical features and outcomes back to the variant. In a genome-first study of patients with TTNtv, individuals of European descent with DCM and TTNtv had increased LV size, decreased LV function, and increased arrhythmia burden compared with patients with DCM without TTNtv, similar to the findings from the previous study. TTNtv were also associated with reduced cardiac function, even in the absence of a DCM.

Figure 4. Proposed workflow for omics-based endotype derivation. (A) Main steps in endotype derivation. The workflow begins by performing high-throughput omics profiling on large cohorts of patients with heart failure (HF). Unsupervised machine learning (ML) methods can then be applied to derive endotypes based on integrated analysis of omics and clinical data. These derived endotypes may be associated with differential outcomes and responses to treatment. Through the process of endotyping we may also identify novel pathways of disease. (B) Endotypes with differential risks of event-free survival. (C) Pathway analysis identifying dysregulated pathways/networks associated with the highest-risk endotypes.
With advances in high-throughput omics technologies, our ability to acquire a multitude of molecular-level data in HF continues to expand. AI techniques are well suited for the integrated analysis of omics and clinical data, and unsupervised ML in particular can uncover novel endotypes of HF. These derived endotypes may be associated with differential outcomes and responses to treatment, with important implications for risk stratification, choice of pharmacotherapy, and clinical trial selection. Through the process of endotyping we may identify novel pathways that are associated with these differential clinical outcomes and treatment responses, thus elucidating new mediators of disease. Pathway analysis may also allow for specification of molecular targets for development of pharmacologic interventions, which will then need to be mechanistically evaluated in experimental model systems (e.g., animal, induced pluripotent stem cells). Proteomics has seen the greatest application of ML for these purposes, but there is potential to apply similar algorithms in other omics fields as well (Figure 4).

One major limitation in these types of studies is that they largely rely on retrospective data. In addition, there is a risk of false discovery in AI-based studies, which is particularly inflated in studies with high-dimensional data and relatively small sample sizes. Successful translation of omics studies to clinical practice will require ongoing studies in large cohorts and validation across diverse populations via global, interdisciplinary collaboration. Pathway and network analysis of multi-omics-level data can also lower the risk of false-positive discovery and confer biological plausibility and interpretability to study findings. Moreover, ML algorithms must be integrated with causal reasoning and clinical knowledge. These integrative approaches will ultimately help to unravel the heterogeneity of this complex syndrome and enable clinicians to individualize care for their patients with HF.

Disclosures / Acknowledgements

None.

Funding

Y.J.S. is supported in part by unrestricted grants from the American Heart Association National Clinical and Population Research Awards, the American Heart Association Career Development Award, Korea Institute of Oriental Medicine, Honjo International Scholarship Foundation, and NIH R01 HL157216.

References


