

CAPÍTULO 9

BIG DATA E O FUTURO DA MEDICINA CARDIOVASCULAR: TRANSFORMANDO DIAGNÓSTICOS E TRATAMENTOS

Gustavo Travassos Gama

RESUMO

O avanço do Big Data tem revolucionado a medicina cardiovascular, proporcionando novas formas de análise e interpretação de dados para a prevenção, diagnóstico e tratamento de doenças cardíacas. Com o aumento exponencial de informações provenientes de registros eletrônicos de saúde, dispositivos vestíveis, exames de imagem e genômica, a aplicação de técnicas avançadas de análise permite identificar padrões, prever riscos e personalizar abordagens terapêuticas. Um dos principais benefícios do Big Data na cardiologia é a capacidade de desenvolver modelos preditivos baseados em aprendizado de máquina. Esses modelos podem avaliar fatores de risco individuais e populacionais, permitindo intervenções precoces e reduzindo complicações cardiovasculares. Além disso, o uso de inteligência artificial e aprendizado profundo aprimora a interpretação de exames como eletrocardiogramas, ressonâncias magnéticas e ecocardiografias, melhorando a acurácia diagnóstica. O monitoramento contínuo de pacientes por meio de dispositivos vestíveis e sensores também é uma inovação impulsionada pelo Big Data. Esses dispositivos coletam informações em tempo real sobre frequência cardíaca, pressão arterial e níveis de atividade física, possibilitando a detecção precoce de arritmias e outros distúrbios cardiovasculares. Isso favorece a personalização do tratamento e a tomada de decisões clínicas mais ágeis e eficazes. No campo da pesquisa, o Big Data contribui para a descoberta de novas abordagens terapêuticas ao integrar informações de diferentes fontes, como ensaios clínicos, bases de dados epidemiológicas e estudos genômicos. Essa integração possibilita a identificação de biomarcadores e o desenvolvimento de terapias personalizadas, alinhadas à medicina de precisão. Apesar dos desafios relacionados à privacidade, interoperabilidade de dados e necessidade de padronização, o Big Data tem um papel essencial no futuro da medicina cardiovascular. Seu uso contínuo pode transformar a assistência médica, tornando-a mais preditiva, personalizada e eficiente, impactando positivamente a saúde da população global.

Palavras-chave: Big Data. Medicina de precisão. Sistemas de informação em saúde. Inteligência artificial. Tecnologia biomédica.

1. O QUE É BIG HEALTH DATA?

Na última década, o termo Big Data tem sido usado para se referir a uma abordagem de pesquisa envolvendo o uso de conjuntos de dados complexos e de larga escala^{1,2}. Embora difícil, pode ser definido como um fenômeno cultural, tecnológico e acadêmico, baseado na aplicação de algoritmos de aprendizado de máquina ao processamento e análise de dados^{3,4}. No campo da biomédica, Big Data é usado para registros eletrônicos de saúde (EHR) considerados relevantes para a compreensão da saúde e da doença, incluindo dados clínicos, de imagem, ômicos, de uso de dispositivos vestíveis, entre outros⁵. O princípio básico é tornar a prática biomédica geradora de evidências, sem a necessidade de projetar e conduzir estudos *ad hoc*^{5,6}. A descrição mais popular para Big Data foi proposta em 2001, por Doug Laney, sendo conhecida no mundo acadêmico como os 5Vs: volume, variedade, velocidade, veracidade e valor⁷.

1.1 Volume

Uma quantidade inimaginável de dados é criada a cada ano e apenas uma parte muito pequena é usada para pesquisa. Para se ter ideia, 1 kilobyte é o tamanho de uma página de texto, 1 gigabyte corresponde a cerca de 6 milhões de livros, e um grande hospital de cuidados terciários gera cerca de 100 terabytes de dados/ano. Isso mostra a possibilidade de armazenar uma enorme quantidade de dados – já que o preço médio do gigabyte de armazenamento caiu nos últimos 30 anos – e no acesso facilitado a eles⁸.

Um alto tamanho de amostra é necessário para investigar tanto a doença cardiovascular rara (DCV) quanto a comum – especialmente quando o desfecho é pouco frequente. Do fim da Segunda Guerra Mundial até o momento, o volume de registros – primeiramente nos Estados Unidos, seguidos pela Europa – tem aumentado gradualmente. O estudo de Framingham começou em 1948 e inscreveu 5.209 pacientes em sua coorte original. Posteriormente, a população aumentou com coortes adicionais – fora da primavera, terceira geração, novo cônjuge, Omni 1 e Omni 2 – e até o momento conta com mais de 15 mil participantes⁹.

Em 2014, por meio de uma inscrição nacional de pacientes com doença cardíaca isquêmica (DCI), o registro SWEDEHEART incluiu 105.674 pacientes com infarto do miocárdio com supradesnívelamento do segmento ST (IAMCSST) e 205.693 com infarto do miocárdio sem supradesnívelamento do segmento ST (IAMSSST)^{10,11}. Outros projetos de pesquisa realizaram uma coleta transnacional de dados, como o estudo PURE¹², que incluiu 140 mil indivíduos de 21 países em cinco continentes entre 2003 e 2018, e o programa CALIBER¹³ para pesquisa de DCV, acessando dados longitudinais de EHR vinculados no Reino Unido, que reuniu uma enorme população de 1,25 milhão de pacientes. Conforme os dados cardiovasculares acumulados crescem, o desejo de analisá-los e convertê-los em resultados também aumenta.

1.2 Variedade

Historicamente, a maioria dos EHR era estruturada em planilhas ou bancos de dados. Porém, a variedade de dados se tornou muito menos congruente e armazenada em inúmeras formas com uma tendência crescente em direção ao formato não estruturado. Dados estruturados são informações altamente organizadas, fáceis de processar e analisar. Exemplos de dados estruturados de DCVs são: idade, doses de medicamentos, valores laboratoriais, eletrocardiograma (ECG), valores ecocardiográficos, dados genéticos – incluindo polimorfismos de nucleotídeo único, variações no número de cópias e mutações raras, ferramentas para análise genética – como matrizes e sequenciamento de próxima geração, entre outros. Dados não estruturados não têm um modelo ou esquema de dados predefinido e podem ser textuais ou não textuais, gerados por humanos ou máquinas¹⁴.

Muitas informações não estruturadas, no entanto, podem ser úteis para montar uma visão holística de um paciente, a exemplo de fatores sociais e ambientais que potencialmente influenciam a saúde. Entre eles, podemos encontrar: instruções médicas, diagnóstico diferencial, relatórios, notas clínicas digitais, exames físicos, mas também blogs, tweets e postagens no Facebook. A maioria dos sistemas de gerenciamento de banco de dados, por meio de mecanismos de intercâmbio e tradução, permite superar as barreiras do passado relacionadas à variedade de formatos de dados incompatíveis, dados não alinhados e semântica de dados inconsistente⁷.

Os obstáculos técnicos na vinculação dessa variedade de informações são um dos principais desafios da análise de Big Data^{14,15}. A ausência de um identificador único de paciente nos Estados Unidos, por exemplo, limitou a vinculação de dados para fins de pesquisa. Mas, o desenvolvimento de algoritmos probabilísticos cada vez mais sofisticados com base nos dados demográficos disponíveis – como nome, idade, código postal, etc. – permite vincular informações com um risco aceitável de erro¹⁴.

1.3 Velocidade

A tecnologia Big Data deve fornecer soluções que reduzam o tempo de armazenamento e recuperação de pacotes de dados – conhecida como latência de dados. A velocidade, de fato, requer arquiteturas que não pressupõem que as informações devem ser quase em tempo real. As empresas desenvolveram soluções, como armazenamentos de dados operacionais, que periodicamente extraem e reorganizam dados para consulta operacional, caches que fornecem acesso instantâneo às informações e roteamento de dados ponto a ponto entre aplicativos e bancos de dados^{7,16}.

Outra possível direção para aumentar a velocidade é a aplicação de algoritmos a qualquer momento, que podem aprender com dados de streaming e que oferecem um resultado valioso se sua execução for interrompida a qualquer momento. A velocidade de criação, armazenamento

e análise de dados é um índice de desempenho da análise de EHR, sendo essencial para sua aplicação na pesquisa cardiovascular. Por outro lado, a alta velocidade na qual os dados são gerados aumentou a lacuna entre o volume de informações disponíveis e a capacidade de analisá-las e interpretá-las^{7,16}.

Veracidade, que corresponde à credibilidade, confiabilidade e precisão dos dados, e valor – dados brutos que, uma vez processados, tornam-se inteligentes, aplicáveis e acionáveis – finalizam a lista de 5Vs, indicando que diferentes canais e fontes podem produzir Big Data, incluindo pesquisas em larga escala, bancos de dados, repositórios e registros (Big Data epidemiológico/clínico), até sequenciamento de última geração e tecnologias de alto rendimento (Big Data molecular) e abordagens computacionais (Big Data infodemiológico ou digital). O Big Data está transformando profundamente as práticas clínicas em práticas disruptivas e informando abordagens orientadas por dados^{7,16}.

2. BIG DATA E CARDIOLOGIA

2.1 Funções e aplicações de Big Data epidemiológico/clínico

Um Big Data epidemiológico/clínico pode ser gerado a partir de pesquisas em larga escala. Esses dados podem informar políticas de saúde pública e global, assim como medicina baseada em evidências e, mais especificamente, cardiologia. Embora os ensaios clínicos controlados randomizados representem o padrão ouro para a construção de um corpo de evidências rigorosas e clinicamente relevantes, eles podem nem sempre refletir populações de pacientes do mundo real, limitando assim a generalização e a validade externa de suas descobertas. Evidências do mundo real, coletadas durante a prática clínica diária, fornecem uma perspectiva complementar a ensaios clínicos controlados rigorosamente e estritamente randomizados^{17,18}. Nesse aspecto, estudos baseados em Big Data são capazes de contribuir para investigações bem projetadas, baseadas em dados pequenos e ensaios clínicos controlados randomizados¹⁸.

Um exemplo de dados do mundo real é o TriNetX – a maior rede global de pesquisa, que fornece evidências do mundo real. Ele concentra dezenas de bilhões de fatos clínicos, diagnósticos, achados laboratoriais, tratamentos recebidos, procedimentos realizados, em mais de 250 milhões de pacientes em todo o mundo, incluindo indivíduos que sofrem de doença hipertensiva, diabetes tipo 2 ou doença renal crônica. No que diz respeito à área cardiológica, essa rede foi explorada para lançar luz sobre o perfil de segurança e os resultados cardiovasculares dos medicamentos¹⁹, a eficácia dos protocolos de reabilitação²⁰, entre outros.

Atualmente, existem muitas fontes geradoras de Big Data epidemiológico, como pesquisas, dados de seguros médicos, dados de registros vitais, dados de coorte, dados de pacientes internados e ambulatoriais, entre outros. Estes dados podem ser recolhidos retrospectiva

ou prospectivamente sendo os registros clínicos prospectivos definidos como conjuntos de dados observacionais grandes/muito grandes, que foram recolhidos prospectiva e sistematicamente e de forma estruturada, para refletir as práticas clínicas do mundo real e os resultados de um determinado procedimento (tratamento ou intervenção cirúrgica) em grandes populações de pacientes, incluindo (sub)populações clínicas/demográficas específicas²¹.

Além disso, além de serem complementares, os ensaios clínicos controlados randomizados podem ser incorporados em registros clínicos. Isso permite economizar tempo e recursos e fortalecer a generalização dos resultados²¹. Por exemplo, o Coronary Artery Surgery Study (CASS) Registry – um dos primeiros exemplos de registros clínicos – é um banco de dados incorporado em um ensaio clínico (a investigação CASS)²².

Na área cardiológica, existem bases de dados e registros clínicos muito grandes, cujas origens podem ser datadas da década de 1980²¹. Os mais populares incluem registros sociais, ou seja, bases de dados endossadas, financiadas e patrocinadas por sociedades científicas, como a Society of Thoracic Surgeons (STS) National Database²³, que coleta resultados clínicos de pacientes submetidos à cirurgia cardiotorácica, e a American Heart Association (AHA) Get With The Guidelines (GWTG) Database, que se baseia em uma iniciativa hospitalar, liderada pela American Heart Association (AHA) e American Stroke Association (ASA), coletando dados de mais de 2.000 hospitais, com o objetivo de melhorar a qualidade do atendimento de pacientes que sofrem de DCV, incluindo insuficiência cardíaca, fibrilação atrial e acidente vascular cerebral²⁴.

Outra importante base de dados social é o ACC National Cardiovascular Data Registry (NCDR), que é composto por 10 registros (oito dos quais são baseados em internamento/procedimentos e os dois restantes em ambulatório), que recolhem dados de mais de 2400 hospitais e 8500 prestadores de cuidados de saúde, superando 60 milhões de registros de pacientes²⁵. Outras bases de dados incluem a Hospital Compare Database, que recolhe dados relativos à qualidade dos cuidados (classificação geral por estrelas e outras medidas de qualidade) de mais de 4.000 hospitais certificados pelo Medicare²⁶, e o Cooperative Cardiovascular Project^{27,28}, um dos primeiros exemplos de um registro clínico.

Alguns conjuntos de dados e registros são dedicados a medicamentos cardiovasculares específicos ou procedimentos cirúrgicos, como o National Heart Lung and Blood Institute (NHLBI) Percutaneous Transluminal Coronary Angioplasty (PTCA) Registry, que concentra dados de resultados para pacientes submetidos a angioplastia coronária transluminal percutânea (ACTP)^{29,30}, o STS/ACC Transcatheter Valve Therapy (TVT) database³¹, que coleta dados de resultados para pacientes submetidos a procedimentos de substituição e reparo de válvulas por cateter de mais 650 locais de notificação, e o registro CathPCI³² do NCDR, que coleta dados de resultados para pacientes submetidos a cateterismo diagnóstico e/ou procedimentos de intervenção coronária percutânea (ICP). Alguns registros e

bases de dados são especificamente dedicados a determinadas doenças cardiovasculares, como o Hypertrophic Cardiomyopathy Registry³³, liderado pela Universidade da Virgínia (Estados Unidos) e pela Universidade de Oxford (Reino Unido), que visa identificar biomarcadores de cardiomiopatia hipertrófica.

Os grandes dados epidemiológicos/clínicos podem ser utilizados para atender a uma variedade de objetivos, incluindo^{34,35};

- realizar uma avaliação epidemiológica de DCV – em termos de incidência, prevalência e taxas de mortalidade;
- quantificar e prever tendências epidemiológicas;
- investigar os determinantes de DCV e comorbidades subjacentes relacionadas;
- identificar marcadores e assinaturas de diagnóstico e prognóstico;
- elaborar ferramentas de pontuação de risco para melhor estratificar pacientes com DCV;
- explorar os resultados clínicos de médio e longo prazo de um determinado procedimento (farmacológico ou cirúrgico) e sua superioridade sobre outro (o concorrente);
- verificar a implementação de recomendações e processos de tomada de decisão, definir referências; e
- calcular os custos econômico-financeiros de uma determinada DCV.

O Big Data pode ajudar a descobrir relações entre doenças e/ou comorbidades, na medida em que tendem a se agrupar, além de contribuir para uma melhor compreensão da chamada medicina de sistema ou rede³⁶. Várias DCVs, incluindo insuficiência cardíaca, frequentemente coexistem com várias comorbidades. Meireles et al³⁷, por exemplo, avaliaram o papel prognóstico e o impacto de várias comorbidades subjacentes no risco de desenvolver insuficiência cardíaca aguda (ICA). Um conjunto de 229 pacientes sofrendo de ICA foi comparado com um conjunto de 201 pacientes com insuficiência cardíaca crônica (ICC). O número de comorbidades foi ligeiramente maior no grupo de pacientes com ICA, incluindo deficiências metabólicas, como hiperuricemia e obesidade, outras DCVs a exemplo de fibrilação atrial (FA) ou doença arterial periférica (DAP), assim como doença renal crônica. Investigar dados como esses auxilia a implementação da cardiologia de precisão ao elaborar intervenções multidimensionais *ad hoc* visando uma subpopulação específica de pacientes.

Ferramentas de risco – que incluem o escore de Framingham, o Global Registry of Acute Coronary Events (GRACE), o Thrombolysis In Myocardial Infarction (TIMI), o Congestive Heart Failure, Hypertension, Age (≥ 75 years), Diabetes, Stroke, Vascular disease, Age (65 to 74 years), and Sex category (CHA2DS2VASc) e o escore de risco Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC)³⁸ – são componentes importantes da chamada cardiologia personalizada, pois permitem estratificar coortes de pacientes e fornecer ao paciente o tratamento de que ele mais precisa.

Exemplos de gerenciamento preciso e personalizado correspondem à avaliação personalizada do fator de risco para uma variedade de DCVs, como fibrilação atrial, isquemia miocárdica crônica, insuficiência cardíaca e hipertensão, dada a constituição biológica individual (genética) e o histórico familiar de doença cardiovascular. Além disso, as disposições farmacológicas, por exemplo, o uso de anticoagulantes, podem ser adaptadas, de forma a minimizar a insurgência de potenciais efeitos colaterais³⁹⁻⁴¹. Porém, existem poucas comparações publicadas entre as diferentes pontuações de risco, o que continua a ser um campo aberto a mais pesquisas e investigações³⁸. Algumas aplicações emergentes de bases de dados apoiadas em Big Data são: abordar as iniquidades e disparidades relacionadas com o sistema cardiovascular, também de uma perspectiva de gênero e realizar análises pós-comercialização de diferentes tratamentos e medicamentos cardiovasculares⁴². O Big Data também é particularmente útil quando a doença cardíaca estudada é rara, como a DCV congênita⁴³.

Considera-se que a inteligência artificial (IA) alavanque e aproveite os bancos de dados baseados em Big Data, superando potencialmente o problema das técnicas estatísticas clássicas e convencionais, a exemplo da análise de pontuação de propensão e da modelagem de regressão multivariada⁴⁴. Ahn et al⁴⁵ desenvolveram o CardioNet, um banco de dados abrangente relacionado a DCV, manualmente curado, padronizado e validado, com base em informações clínicas (estruturadas ou não estruturadas) coletadas de 748.474 pacientes, que pode ser utilizado para análises de IA e fornecer insights sobre o tratamento de pacientes com DCV.

Barbieri et al⁴⁶ combinaram a análise de sobrevivência clássica – modelagem de risco proporcional de Cox – com uma abordagem de aprendizado profundo em uma coorte de 2.164.872 neozelandeses com idades entre 30 e 74 anos. Os preditores de eventos de DCV foram o uso de tabaco em mulheres e a doença pulmonar obstrutiva crônica (DPOC) com infecção respiratória inferior aguda em homens, além de fatores de risco bem estabelecidos, como pressão alta, dor no peito, diabetes e comprometimento metabólico.

Apesar do uso de ferramentas estatísticas sofisticadas, ainda existem questões em aberto que precisam ser abordadas. Estudos baseados em Big Data podem oferecer um ponto de vista diferente, mas algumas descobertas conflitantes de estudos clínicos randomizados controlados e investigações pequenas e bem conduzidas podem ser encontradas. Tais discrepâncias podem ser devidas à natureza única do banco de dados usado no estudo: cada banco de dados cardiológico varia significativamente nos métodos empregados para coletar e capturar dados e na(s) população(ões) que ele representa especificamente¹⁸.

Além disso, o formato do banco de dados (estruturado vs. não estruturado) pode impactar a qualidade dos dados. Hernandez-Boussard et al¹⁸, por exemplo, extraíram um conjunto de dados que incluía 10.840 notas clínicas e encontraram menores taxas de recall e precisão (51,7 e 98,3%,

respectivamente) no caso de prontuários eletrônicos de saúde (EHR) estruturados, em relação a EHR não estruturados (95,5 e 95,3%, respectivamente), garantindo a medição rotineira de recall para cada banco de dados/registo, antes de prosseguir com o processamento e análise de dados.

Repositórios, registros e bancos de dados de Big Data, portanto, são cada vez mais comuns no campo da prática cardiológica e pesquisa clínica. Existem, no entanto, variações consideráveis significativas em características sociodemográficas, comorbidades e principais taxas de complicações entre estudos individuais (unicêntricos ou multicêntricos), baseados em banco de dados, e até mesmo entre os próprios estudos de registro, como banco de dados clínico vs. administrativo. Isso deve ser levado em conta ao avaliar criticamente a pesquisa cardiológica e a modelagem de ajuste de risco²¹.

As bases de dados administrativas, especificamente, podem fornecer muitas informações sobre epidemiologia de doenças, comorbidades, disparidades e desigualdades no acesso a cuidados de saúde e resultados clínicos, além de informar de forma orientada por dados os processos de tomada de decisão subjacentes aos tratamentos farmacológicos cardiológicos ou procedimentos cirúrgicos, em termos de parâmetros de estratificação de risco pré-operatório para reduzir/minimizar as taxas de morbidade e mortalidade perioperatórias. Por outro lado, as bases de dados administrativas podem sofrer imprecisões administrativas, viés de registro, mudanças temporais nos sistemas de nosologia e nomenclatura, códigos de cobrança, assim como uma escassez de parâmetros clinicamente relevantes, incluindo variáveis e resultados específicos da cardiologia²¹.

Uma questão importante que limita a implantação de bancos de dados e registros está relacionada à sua interoperabilidade e, às vezes, ao uso inconsistente de definições. Nem todos os bancos de dados também atendem aos padrões regulatórios e são suficientemente curados/validados. Nesse caso, a padronização de dados e os metadados são urgentemente garantidos^{18,21}.

Embora bem projetados e conduzidos, estudos clínicos, especialmente aqueles que dependem de Small Data, muitas vezes, são estatisticamente subpotentes e comprometidos por vários vieses, incluindo viés de amostragem e seleção de participantes, o que dificulta a generalização dos achados, com amostras não representativas de toda a população. Também é difícil estratificar de acordo com um determinado tratamento farmacológico cardiológico ou procedimento cirúrgico se a amostra for muito heterogênea e seu tamanho impedir a realização de cálculos estatisticamente robustos e confiáveis. A confiança e a certeza podem aumentar com Big Data, paralelamente, no entanto, ao crescimento da complexidade e dos custos computacionais associados^{47,48}.

Bancos de dados baseados em Big Data também podem ser afetados por vieses, como vieses de registro ou associação e outros artefatos estatísticos, como epidemiologia reversa ou causalidade reversa. Alguns

estudos baseados em banco de dados, por exemplo, descobriram que o índice de massa corporal (IMC), o perfil lipídico e a pressão arterial, que geralmente preveem um resultado clínico ruim na população em geral, se tornam preditores de prognóstico inverso em pacientes com insuficiência cardíaca crônica. Uma maior sobrevivência foi associada ao sobrepeso e obesidade, hipercolesterolemia e altos valores de pressão arterial, o que é bastante contraintuitivo. Várias hipóteses foram formuladas, incluindo a presença de síndrome do complexo desnutrição-inflamação ou síndrome desnutrição-inflamação-caquexia. Porém, alguns pesquisadores acreditam que é mais provável (e biologicamente/clinicamente plausível) que essas descobertas sejam artefatos estatísticos⁴⁹.

2.2 Funções e aplicações do Big Data molecular

Tecnologias de laboratório de alto rendimento, incluindo chips de microarray, sequenciamento de DNA e RNA de última geração e sequenciamento de exoma completo, sequenciamento acoplado à imunoprecipitação de cromatina, análise de interação de cromatina por sequenciamento de marcadores de extremidade pareada (ChIA-PET), captura de conformação de cromatina com sequenciamento, ensaio para cromatina acessível por transposase com sequenciamento de alto rendimento (ATAC-Seq) e análise proteômica baseada em espectrometria de massa podem aprimorar o Big Data molecular, abrindo caminho para uma cardiologia personalizada/individualizada^{50,51}.

Os Big Data moleculares podem elucidar os mecanismos subjacentes à etiopatogenia de uma determinada doença cardíaca e identificar novos alvos potencialmente medicamentáveis para o desenvolvimento de terapias farmacológicas *ad hoc*. A cardiologia personalizada pode se beneficiar de estudos de associação genômica ampla e pós-genômica, visando à identificação de novos fatores de transcrição cardiogênicos, validações genótípicas e fenótípicas de potenciais reguladores transcricionais e mecanismos moleculares/celulares^{52,53}. A CardioGenBase⁵² é uma ferramenta de recursos on-line abrangente e baseada na literatura, que coleta associações gene-doença (mais de 1.500) para as principais DCVs, incluindo doença cerebrovascular, doença cardíaca isquêmica, doença arterial coronária (DAC), doença cardíaca inflamatória, doença cardíaca reumática, doença cardíaca hipertensiva, entre outras.

Vakili et al⁵³ fizeram esforços para combinar todas as especialidades baseadas em OMICS dentro de uma abordagem multi-OMICS altamente integrada e coerente denominada panomics, para elucidar a patogênese multifatorial da DCV. Os autores exploraram sistematicamente a literatura e conseguiram encontrar 104 bancos de dados baseados em OMICS relacionados à DCV, com 72 fornecendo genômica/pós-genômica e medições clínicas. Destes conjuntos de dados, 59 e 65 bancos de dados eram transcriptômicos, epigenômicos/metilômicos, 41 proteômicos, 42 metabolômicos e 22 microbiômicos.

Combinando literatura acadêmica, informações clínicas e baseadas em OMICS, e explorando a abordagem doença, Sarajlić et al⁵⁴ avaliaram a estrutura da rede de interação proteína-proteína humana para descobrir novos genes relacionados a DCVs, que poderiam ser alvos potencialmente tratáveis. Os autores descobriram que esses novos genes estavam envolvidos em cascatas de sinalização intracelular, atividade de transdução de sinalização, ligação enzimática e vias de sinalização mediadas por receptores intracelulares.

Além disso, a convergência única entre diferentes disciplinas, como nano-(bio)-engenharia, impressão tridimensional (3D) e simulação computacional, modelagem molecular e matemática, e técnicas bioestatísticas avançadas e sofisticadas e IA (data mining, machine e deep learning), está moldando novos caminhos e criando oportunidades no campo da prática cardiológica e pesquisa clínica, tornando-as mais multi e interdisciplinares e complexas, e mais capazes de abordar os desafios biomédicos^{55,56}.

A impressão 3D, por sua vez, está sendo cada vez mais utilizada na biomedicina e, em particular, na cardiologia. Geralmente, são produzidos principalmente modelos anatômicos rígidos, mas considera-se que a incorporação da funcionalidade dinâmica avance dramaticamente no planejamento cirúrgico cardiovascular pré-operatório, assim como na hemodinâmica. Os modelos 3D podem esclarecer diferentes condições fisiopatológicas relacionadas à DCV, complementando assim as informações obtidas usando imagens clássicas⁵⁷. O Big Data molecular, sozinho ou combinado/integrado ao Big Data epidemiológico, também pode capturar o panorama de diversas doenças e eventos cardiológicos, idiopáticos ou congênitos, incluindo cardiomiopatia dilatada e insuficiência cardíaca, entre outros^{58,59}.

2.3 Funções e aplicações de Big Data gerados por técnicas de imagem, tecnologias vestíveis e sensores inteligentes

As mais recentes conquistas tecnológicas no campo da saúde móvel (mHealth) e da saúde ubíqua (uHealth), com smartphones, dispositivos inteligentes, smartwatches e outros sensores vestíveis estão revolucionando o campo da cardiologia, envolvendo diretamente o paciente, melhorando a sua adesão e conformidade terapêuticas e também permitindo sua monitorização remota. Sensores vestíveis de diferentes tipos – como dispositivos vestíveis bioelétricos, mecanoelétricos, optoeletrônicos e ultrassônicos – facilitam a coleta contínua de sinais vitais cardiovasculares, incluindo pressão arterial, frequência cardíaca e ritmo cardíaco, saturação de oxigênio no sangue e glicemia, além de ondas cerebrais, qualidade do ar, exposição a radiações e outras métricas, e auxiliam na intervenção precoce^{60,61}.

Gandhi et al⁶² investigaram a eficácia das intervenções mHealth para a prevenção secundária de DCV, reunindo 27 estudos, com um total de 5.165

pacientes. Eles identificaram que o mHealth aumenta a adesão terapêutica, assim como a conformidade geral, farmacológica ou não farmacológica. Diferentes metas tiveram maior probabilidade de serem atingidas, correspondentes a pressão arterial, exercício e atividade física com redução do tempo sedentário e sentado, porém sem cessação do tabagismo e redução do perfil lipídico. No entanto, o grupo mHealth não diferiu do grupo de tratamento padrão em termos de hospitalizações e readmissões hospitalares. Poucos estudos mostraram uma redução significativa na angina e na recorrência de ataque isquêmico transitório/AVC em pacientes com doença cerebrovascular. A taxa de mortalidade cardiovascular foi computada como sendo menor, mesmo não atingindo o limiar de significância.

Wali et al⁶³ destacaram que as intervenções de mHealth podem ser particularmente úteis para alcançar comunidades vulneráveis e carentes, incluindo indivíduos aborígenes e indígenas ou sujeitos que residem em países de baixa e média renda. Normalmente, esses indivíduos são excluídos ou são sub-representados em ensaios clínicos.

Gamificação e aplicativos móveis gamificados (apps) representam outra ramificação interessante e promissora da área da saúde digital. Em revisão sistemática da literatura, Davis et al⁶⁴ sintetizaram sete estudos, totalizando 657 pacientes. Eles descobriram que a gamificação resultou em melhor adoção de estilos de vida e comportamentos mais saudáveis, como, por exemplo, em termos de prática de exercícios e atividade física, além de melhor perfil bioquímico, humor e motivação aprimorados. A alfabetização e o conhecimento em saúde relacionados a DCV melhoraram de forma significativa, embora alguns parâmetros, como pressão arterial, índice de massa corporal, autogerenciamento e conformidade terapêutica, fossem comparáveis ao padrão de atendimento.

Em relação aos dispositivos inteligentes, como smartwatches e smartphones, Prasitlunkum et al⁶⁵ avaliaram quantitativamente a precisão da utilização de dispositivos vestíveis para triagem, detecção e diagnóstico adequado da fibrilação atrial. Os autores foram capazes de calcular áreas excelentes sob as curvas de características operacionais do receptor resumidas (SROC) em 0,96 e 0,94, para smartphones e smartwatches, respectivamente. A sensibilidade e a especificidade estavam na faixa de 94 a 96% e 93 a 94% para os dois tipos de dispositivos inteligentes/vestíveis, respectivamente. Eles provaram ser tão precisos e confiáveis em termos de diagnóstico quanto os padrões ouro, como a fotopletoxiografia e a eletrocardiografia de derivação única.

Sinais e dados gerados por técnicas de imagem, como eletrocardiografia, tomografia computadorizada ou ressonância magnética, podem ser processados, analisados e refinados usando IA⁶⁶. Por exemplo, a MOCOnet é uma rede neural convolucional de última geração, que pode melhorar e aprimorar o mapeamento quantitativo de T1 por ressonância magnética cardiovascular, tornando-o mais robusto, confiável, clinicamente significativo, menos propenso a artefatos de movimento e de maneira

eficiente em termos de tempo. A MOCOnet, sendo puramente orientada por dados, supera os métodos atualmente disponíveis para correção de movimento, que são orientados por modelo⁶⁷.

A radiômica e a radiogenômica, por sua vez, são campos de pesquisa translacionais altamente inovadores, que visam à mineração, recuperação, fusão, processamento, análise e extração de padrões e interpretações clinicamente significativos de conjuntos de dados de grande escala e alta dimensão gerados por técnicas e ferramentas de imagem clínica, incluindo angiografia por tomografia computadorizada cardíaca e ressonância magnética cardíaca⁶⁸.

Os últimos avanços relativos a protocolos cada vez mais sofisticados permitem integrar recursos de imagem e perfis moleculares para identificar biomarcadores e assinaturas relevantes e clinicamente significativos – como lesões ateroscleróticas, placas coronárias e anormalidades estruturais do miocárdio – relacionados ao diagnóstico, prognóstico e resposta ao tratamento. A IA supervisionada e não supervisionada, incluindo aprendizado profundo e de máquina, pode combinar e agregar ainda mais dados e auxiliar no desenvolvimento de modelos de risco e ferramentas que podem facilitar procedimentos clínicos de diagnóstico e prognóstico⁶⁸.

No campo da pesquisa cardiológica, a radiômica e a radiogenômica podem ser utilizadas para a caracterização, criação de perfil/fenotipagem e estratificação de risco de doença cardíaca coronária, cardiomiopatia hipertrófica, doença cardíaca isquêmica e doença cerebrovascular, entre outras. Dada sua atualidade, ainda há muito a ser explorado neste campo. Por outro lado, pode-se antecipar que a radiômica, a radiogenômica e outros Big Data gerados por dispositivos e sensores vestíveis/inteligentes impactarão profundamente tanto a prática quanto a pesquisa cardiológica⁶⁸⁻⁷⁰.

2.4 Funções e aplicações do Big Data infodemiológico

A infodemiologia – junção de informação e epidemiologia – e a infovigilância – combinação de informação e vigilância – representam uma disciplina altamente inovadora, na intersecção da ciência da computação, dos dados e do comportamento, com o objetivo avaliar os determinantes das atividades computacionais e digitais, como consultas na web, uso de mídias sociais, publicação em redes sociais e produção/consumo de material on-line. Pesquisadores da área de infodemiologia e infovigilância utilizam recursos que permitem avaliar a demanda e o consumo de informação, como o Google Trends, uma ferramenta de código aberto que permite rastrear e monitorar pesquisas na web realizadas pelo mecanismo de busca Google^{71,72}.

Tanto a infodemiologia como a infovigilância permitem rastrear a eficácia de campanhas de conscientização, como a Go Red for Women, uma iniciativa social que visa melhorar e aumentar a alfabetização relacionada a DCV e AVC entre mulheres⁷³. Suero-Abreu et al⁷⁴ investigaram o impacto dessa campanha no comportamento de busca de informações sobre saúde,

utilizando o Google Trends, encontrando volumes de pesquisa aumentados e vários termos relacionados a DCV ao longo de 15 anos. Porém, não foi encontrado aumento nas pesquisas digitais relacionadas a AVC durante o período do estudo.

Dzaye et al⁷⁵ também exploraram técnicas de infodemiologia e infovigilância para avaliar o interesse público em relação à DCV e comorbidades relacionadas durante a pandemia de COVID-19. De acordo com alguns estudos, a atenção à DCV teria diminuído, apesar da relação negativa entre DCV e infecção. Pacientes que sofrem de DCV ou com fatores de risco para DCV têm sido consistentemente relatados como apresentando resultados piores do que seus equivalentes sem DCV. Os autores descobriram que o interesse digital em termos como exercício ou atividade física e cigarros aumentou (em 18%) e diminuiu (em 52,5%), respectivamente. Mas, o interesse em termos como estatina, perfil lipídico, lipoproteína de baixa densidade (LDL) e hemoglobina A1C aumentou significativamente após um declínio anterior ao longo do tempo.

Por outro lado, de acordo com outro estudo de pesquisa realizado pelo mesmo grupo, os primeiros meses da pandemia de COVID-19 foram acompanhados por uma diminuição no interesse de pesquisa por infarto do miocárdio e síndrome coronariana aguda (SCA), explicando potencialmente o excesso de mortalidade cardiovascular, apesar de uma redução acentuada na hospitalização por SCA⁷⁶. Esses resultados indicam que mecanismos de busca e outros fluxos de dados não convencionais parecem ser ferramentas valiosas e promissoras, capazes de fornecer insights sobre comportamentos de busca de informações sobre saúde e avaliar a eficácia de campanhas sociais e outras intervenções^{75,76}.

A qualidade dos sites relacionados à cardiologia e dos materiais disponíveis on-line é altamente heterogênea e variável tanto em termos de conteúdo quanto de informações fornecidas. Azer et al⁷⁷, por exemplo, avaliaram a qualidade, precisão e legibilidade das páginas da Wikipédia sobre DCV. Cerca de 83% das páginas da Wikipédia foram consideradas de qualidade moderada, com 8,5% sendo de boa e má qualidade, respectivamente. Apesar da apresentação clínica e etiopatogenia da DCV serem tratadas e discutidas, várias seções, incluindo a fisiopatologia, sinais e sintomas, diagnóstico e tratamento, nem sempre foram precisas e adequadamente referenciadas academicamente. Várias entradas exibiram erros e omissões. A legibilidade estava no nível de disciplinas universitárias.

Pacientes com DVC usam a internet como uma fonte de baixo custo e facilmente disponível de informações pessoais sobre saúde, para aprender mais sobre sua condição/distúrbio, bem como sobre potenciais opções de tratamento e médicos e cirurgias de DVC. Segundo uma pesquisa de Jones et al⁷⁸, 74,3% dos entrevistados navegavam na internet, com 63% utilizando-a diariamente. No caso de o paciente não conseguir acessar a web diretamente, um membro da família estava disposto a fazê-lo em seu nome.

Os autores concluíram que a maioria dos pacientes (cerca de 85%) utilizou a internet, estando particularmente interessado em informações locais.

3. BIG DATA E CARDIOLOGIA PARTICIPATIVA

O Big Data também pode contribuir para uma superespecialização emergente no campo da cardiologia: a chamada cardiologia participativa, promovendo assim a participação pública no campo da prática cardiológica e da pesquisa clínica, com a criação de redes sociais colaborativas globais e integração da cardiologia básica e translacional com a ciência cidadã. Isso é de suma importância, especialmente em países de baixa e média renda, e ajudaria a conter/reduzir as disparidades e iniquidades em saúde^{79,80}. Em estudo, Wali et al⁶³ constataram que o estabelecimento de parcerias e relacionamentos colaborativos com membros da comunidade – especialmente aqueles de populações carentes e vulneráveis – melhoraria e aumentaria significativamente a eficácia da intervenção cardiológica, garantindo que ela fosse concebida e implementada dentro do contexto apropriado.

A cardiologia participativa, como um ramo da medicina participativa, dá um novo valor e importância ao paciente. Os últimos avanços científicos e tecnológicos e as tendências atuais na prática clínica e na pesquisa, especialmente na área cardiológica, mudaram gradualmente a atenção e o interesse dos profissionais para os resultados subjetivos do paciente (satisfação, dor, qualidade de vida, etc.), além dos resultados clínicos objetivos (absorção e consumo de recursos de saúde, processos de saúde e prestação de serviços, taxas de morbidade e mortalidade)⁸¹.

Porém, para a maioria dos cardiologistas e cirurgiões cardiológicos, o mundo do Programmable Read-Only Memory (PROM) representa um ambiente de assistência médica ainda desconhecido, cuja navegação, ao incorporar missão, valores e cultura, pode promover a prática e a pesquisa cardiológicas. Existem várias lacunas na implementação e incorporação completa dos PROMs na prática cardiológica de rotina diária⁸¹⁻⁸³.

De acordo com o International Consortium for Health Outcomes Measurement (ICHOM), embora existam vários registros nacionais, internacionais e de ensaios para insuficiência cardíaca, poucos podem ser considerados centrados no paciente e faltam diretrizes e listas de verificação padronizadas que orientem o processo de uso adequado, eficaz e significativo dos PROMs. Para preencher essa lacuna, o ICHOM desenvolveu um conjunto de dados de 17 itens, que consiste em vários domínios: resultados funcionais, psicossociais, de carga de cuidados e relacionados à sobrevivência. Este conjunto, que também inclui PROMs além de medições clínicas/objetivas e dados administrativos, permite comparar o tratamento e o gerenciamento da insuficiência cardíaca entre vários provedores de saúde e várias regiões, globalmente⁸⁴.

4. OUTRAS LIMITAÇÕES E OPORTUNIDADES NO USO DE BIG DATA PARA O FUTURO DA MEDICINA CARDIOVASCULAR

O EHR é reconhecido como um repositório de diversos conjuntos de dados longitudinais, que informam sobre DCVs e incorporam dados laboratoriais, de imagem, biológicos e descritivos com heterogeneidade na frequência de coleta de dados. Sendo assim, o EHR está preparado para o aprimoramento de dados⁸⁵.

Apesar da importância de suas informações, o EHR carece de padronização e uma linguagem universal global para facilitar a interoperabilidade entre pesquisadores e em diferentes plataformas. A capacidade de coletar todas as variáveis de dados, incluindo aquelas que não se considera estarem associadas ao (pató)fenótipo de interesse (características “ortogonais”), o EHR também depende da profundidade das ferramentas analíticas disponíveis e está sujeita aos mesmos problemas de qualidade, confiabilidade e inconsistência de dados observados em qualquer fonte de dados em larga escala⁸⁶.

Foram introduzidas metodologias para validar os EHRs propensos a erros, sendo provável que elas evoluam à medida que as fontes de dados se tornam cada vez maiores e mais complexas. Assim, a assimilação de dados de EHR na fenotipagem de precisão e no tratamento cardiovascular continua sendo um desafio fundamental e contínuo⁸⁷.

O aumento do tamanho dos conjuntos de dados de EHR torna ainda mais necessária a utilização de computação de alto desempenho (supercomputação com processamento paralelo) com uma mudança para a computação exascale: um sistema computacional que pode executar um quíntilhão de 10^{18} cálculos por segundo^{88,89}. Essa computação de alto desempenho irá acelerar e facilitar o uso da aprendizagem de máquina e da IA na medicina clínica em geral e na medicina cardiovascular em particular. Entre suas muitas aplicações úteis de uma perspectiva de Big Data, a IA informou análises de imagens médicas cardiovasculares e melhorou a fenotipagem; tomada de decisão clínica e previsão de risco; identificação de novos grupos ou coortes de fenótipos; e análises genômicas-fenômicas de conjuntos de dados complexos nos quais relações anteriormente desconhecidas podem ser reveladas⁹⁰⁻⁹².

Embora ensaios clínicos de referência tenham informado a endofenotipagem e a fenotipagem em doenças cardiovasculares, a demonstração contínua da heterogeneidade da doença exige que a relação genótipo–endofenótipo–fenótipo clínico, por exemplo, seja continuamente refinada. Os recursos disponíveis para melhorar a fenotipagem de precisão, no entanto, são limitados. Assim, estudos históricos e conjuntos de dados podem ter aplicabilidade limitada para análises futuras. Embora a análise de Big Data possa superar essa limitação, compilando dados díspares de um coletivo de amostra de grande porte, esse desafio também mostra a necessidade de considerar a relevância futura dos dados coletados em

ensaios clínicos e a criação de estudos inclusivos e integrados em larga escala em um nível (multi)nacional⁹⁰⁻⁹².

Entre as oportunidades, redes clínicas são relatadas em pequena escala, indicando que uma abordagem modificada usando redes correlacionais já provou ser útil em vários locais, incluindo esquemas baseados em comorbidade, que melhoram o diagnóstico de DCVs pediátricas e relatórios que esclarecem a heterogeneidade clínica de pacientes com doença pulmonar obstrutiva crônica, uma comorbidade comum de DCV^{93,94}.

Essa abordagem pode ser útil para decifrar resultados de testes de diagnóstico cardiovascular, que dependem da lógica da cadeia de ramificação para organizar dados. Esses métodos podem ignorar variáveis inter-relacionadas que informam patofisiologias diferenciadas ou são eficazes para identificar subgrupos de pacientes. Por exemplo, o teste de exercício cardiopulmonar invasivo é um teste abrangente usado para diagnosticar dispneia inexplicada e gera aproximadamente 100 medições por paciente, com base em sete parâmetros fisiológicos, incluindo hemodinâmica cardiopulmonar central e troca gasosa respiratória. Os métodos atuais usados para interpretar esses testes, no entanto, geralmente se concentram em um subconjunto muito pequeno de informações, geralmente menos de cinco variáveis⁹⁵.

Uma rede de exercícios, que incluía 39 nós e 98 links, foi relatada recentemente fornecendo informações sobre relações inesperadas entre medições de testes, correspondentes a variáveis em vários parâmetros de exercícios diferentes, como função pulmonar, função ventricular direita. Essa rede foi reduzida ainda mais a um grupo de 10 variáveis, o que, por sua vez, foi eficaz para identificar quatro subgrupos distintos de pacientes, definidos por perfis clínicos, de exercícios e de resultados exclusivos. A partir dessa abordagem, foi montado um modelo de predição de risco baseado na medicina de rede, que se mostrou superior aos métodos de regressão linear probabilística (tradicional) para estratificação de risco⁹⁶.

Essas e outras novas abordagens para classificar pacientes, que são baseadas principalmente em uma coleção de parâmetros clínicos relacionados, podem ser importantes para a seleção de fenótipos em abordagens de pensamento avançado para o desenho de ensaios clínicos de medicina cardiovascular. Esses recursos incluem critérios de inscrição baseados na randomização mendeliana, loci de características quantitativas e desenhos de ensaios adaptativos usando métodos baseados em farmacologia de sistemas que permitem flexibilidade na inscrição de pacientes, cronograma de coleta de dados e seleção de endpoints combinados com a coleta contínua de dados em frentes clínicas e farmacológicas⁹⁷.

REFERÊNCIAS

1. Science Staff. Dealing with data. Challenges and opportunities. Introduction. Science. 2011;331:692–3.
2. Denaxas SC, Morley KI. Big biomedical data and cardiovascular disease research: opportunities and challenges. Eur Heart J Qual Care Clin Outcomes. 2015;1:9–16.
3. Boyd D, Crawford K. Critical questions for big data. Inform Commun Soc. 2012;15:662–79.
4. Sim I. Two ways of knowing: Big data and evidence-based medicine. Ann Intern Med. 2016;164:562–3.
5. Hemingway H, Asselbergs FW, Danesh J, Dobson R, Maniadakis N, Maggioni A, et al. Big data from electronic health records for early and late translational cardiovascular research: challenges and potential. Eur Heart J. 2018;39:1481–95.
6. Attisano T, Silverio A, Stabile E, Briguori C, Tuccillo B, Scotto Di Uccio F, et al. Safety and feasibility of balloon aortic valvuloplasty in non-TAVI centers: the “BAV for life” experience. Catheter Cardiovasc Interv. 2018;93:E63–70.
7. Laney D. 3D data management: controlling data volume, velocity, and variety. Meta Group. 2001.
8. Austin C, Kusumoto F. The application of big data in medicine: current implications and future directions. J Interv Card Electrophysiol. 2016;47:51–9.
9. Mahmood SS, Levy D, Vasan RS, Wang TJ. The Framingham Heart Study and the epidemiology of cardiovascular disease: a historical perspective. Lancet. 2014;383:999–1008.
10. Szummer K, Wallentin L, Lindhagen L, Alfredsson J, Erlinge D, Held C, et al. Improved outcomes in patients with ST-elevation myocardial infarction during the last 20 years are related to implementation of evidence-based treatments: experiences from the SWEDEHEART registry 1995-2014. Eur Heart J. 2017;38:3056–65.
11. Szummer K, Wallentin L, Lindhagen L, Alfredsson J, Erlinge D, Held C, et al. Relations between implementation of new treatments and improved outcomes in patients with non-ST-elevation myocardial infarction during the last 20 years: experiences from SWEDEHEART registry 1995 to 2014. Eur Heart J. 2018;39:3766–76.

12. Dehghan M, Mente A, Rangarajan S, Sheridan P, Mohan V, Iqbal R, et al. Association of dairy intake with cardiovascular disease and mortality in 21 countries from five continents (PURE): a prospective cohort study. *Lancet*. 2018;392:2288–97.
13. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1.25 million people. *Lancet*. 2014;383:1899–911.
14. Weber GM, Mandl KD, Kohane IS. Finding the missing link for big biomedical data. *JAMA*. 2014;311:2479–80.
15. Wang Z, Shah AD, Tate AR, Denaxas S, Shawe-Taylor J, Hemingway H. Extracting diagnoses and investigation results from unstructured text in electronic health records by semi-supervised machine learning. *PLoS One*. 2012;7:e30412.
16. Gligorijevic V, Malod-Dognin N, Przulj N. Integrative methods for analyzing big data in precision medicine. *Proteomics*. 2016;16:741–58.
17. Hemkens LG. How routinely collected data for randomized trials provide long-term randomized real-world evidence. *JAMA Netw Open*. 2018;1:e186014.
18. Hernandez-Boussard T, Monda KL, Crespo BC, Riskin D. Real world evidence in cardiovascular medicine: ensuring data validity in electronic health record-based studies. *J Am Med Inform Assoc*. 2019;26:1189–94.
19. Stapff M, Hilderbrand S. First-line treatment of essential hypertension: a real-world analysis across four antihypertensive treatment classes. *J Clin Hypertens (Greenwich)*. 2019;21:627–34.
20. Buckley BJR, Harrison SL, Fazio-Eynullayeva E, Underhill P, Sankaranarayanan R, Wright DJ, et al. Cardiac rehabilitation and all-cause mortality in patients with heart failure: a retrospective cohort study. *Eur J Prev Cardiol*. 2021;28:1704–10.
21. Meltzer SN, Weintraub WS. The role of national registries in improving quality of care and outcomes for cardiovascular disease. *Methodist Debakey Cardiovasc J*. 2020;16:205–11.

22. Myers WO, Blackstone EH, Davis K, Foster ED, Kaiser GC, CASS. Registry long-term surgical survival. Coronary artery surgery study. *J Am Coll Cardiol.* 1999;33:488–98.
23. Jacobs JP, Shahian DM, Prager RL, Edwards FH, McDonald D, Han JM, et al. Introduction to the STS national database series: outcomes analysis, quality improvement, and patient safety. *Ann Thorac Surg.* 2015;100:1992–2000.
24. Smaha LA. American Heart Association. The American Heart Association Get With The Guidelines program. *Am Heart J.* 2004;148(5 Suppl):S46–8.
25. Brindis RG, Fitzgerald S, Anderson HV, Shaw RE, Weintraub WS, Williams JF. The American College of Cardiology-National Cardiovascular Data Registry (ACC-NCDR): building a national clinical data repository. *J Am Coll Cardiol.* 2001;37:2240–5.
26. Shwartz M, Ren J, Peköz EA, Wang X, Cohen AB, Restuccia JD. Estimating a composite measure of hospital quality from the hospital compare database: differences when using a Bayesian hierarchical latent variable model vs. denominator-based weights. *Med Care.* 2008;46:778–85.
27. Gunter N, Moore L, Odom P. Cooperative cardiovascular project. *J S C Med Assoc.* 1997;93:177–9.
28. Ramunno LD, Dodds TA, Traven ND. Cooperative cardiovascular project (CCP) quality improvement in Maine, New Hampshire, and Vermont. *Eval Health Prof.* 1998;21:442–60.
29. Kelsey SF, Miller DP, Holubkov R, Lu AS, Cowley MJ, Faxon DP, et al. Results of percutaneous transluminal coronary angioplasty in patients ≥65 years (from the 1985 to 1986 national heart, lung, and blood institute's coronary angioplasty registry). *Am J Cardiol.* 1990;66:1033–8.
30. Ayalon N, Jacobs AK. Incomplete revascularization in patients treated with percutaneous coronary intervention: when enough is enough. *JACC Cardiovasc Interv.* 2016;9:216–8.
31. Hansen JW, Foy A, Yadav P, Gilchrist IC, Kozak M, Stebbins A, et al. Death and dialysis after transcatheter aortic valve replacement: an analysis of the STS/ACC TVT registry. *JACC Cardiovasc Interv.* 2017;10(20):2064–75.
32. Moussa I, Hermann A, Messenger JC, Dehmer GJ, Weaver WD, Rumsfeld JS, et al. The NCDR CathPCI registry: a US national perspective on care and

outcomes for percutaneous coronary intervention. *Heart*. 2013;99(4):297–303.

33. Kramer CM, Appelbaum E, Desai MY, Desvigne-Nickens P, DiMarco JP, Friedrich MG, et al. Hypertrophic cardiomyopathy registry: the rationale and design of an international, observational study of hypertrophic cardiomyopathy. *Am Heart J*. 2015;170(2):223–30.

34. Xiao H, Ali S, Zhang Z, Sarfraz MS, Zhang F, Faisal M. Big data, extracting insights, comprehension, and analytics in cardiology: an overview. *J Healthc Eng*. 2021;2021:6635463.

35. Shah RU, Rumsfeld JS. Big Data in Cardiology. *Eur Heart J*. 2017;38(25):1865–7. doi: 10.1093/eurheartj/ehx284.

36. Barabási AL. Network medicine—from obesity to the “diseasome”. *N Engl J Med*. 2007;357(4):404–7.

37. Meireles MA, Golçalves J, Neves J. Acute heart failure comorbidome: the impact of everything else. *Acta Med Port*. 2020;33(2):109–15.

38. Cohn SL, Fernandez Ros N. Comparison of 4 cardiac risk calculators in predicting postoperative cardiac complications after non-cardiac operations. *Am J Cardiol*. 2018;121(1):125–30.

39. Lee MS, Flammer AJ, Lerman LO, Lerman A. Personalized medicine in cardiovascular diseases. *Korean Circ J*. 2012;42(8):583–91.

40. Auffray C, Charron D, Hood L. Predictive, preventive, personalized and participatory medicine: back to the future. *Genome Med*. 2010;2(1):57.

41. Sagner M, McNeil A, Puska P, Auffray C, Price ND, Hood L, et al. The P4 health spectrum - a predictive, preventive, personalized and participatory continuum for promoting healthspan. *Prog Cardiovasc Dis*. 2017;59(5):506–21.

42. Lee KT, Hour AL, Shia BC, Chu PH. The application and future of big database studies in cardiology: a single-center experience. *Acta Cardiol Sin*. 2017;33(6):581–7.

43. Van den Eynde J, Manlihot C, Van De Bruaene A, Diller GP, Frangi AF, Budts W, et al. Medicine-based evidence in congenital heart disease: how artificial intelligence can guide treatment decisions for individual patients. *Front Cardiovasc Med*. 2021;8:798215.

44. de Marvao A, Dawes TJ, Howard JP, O'Regan DP. Artificial intelligence and the cardiologist: what you need to know for 2020. *Heart*. 2020;106(5):399–400.
45. Ahn I, Na W, Kwon O, Yang DH, Park GM, Gwon H, et al. CardioNet: a manually curated database for artificial intelligence-based research on cardiovascular diseases. *BMC Med Inform Decis Mak*. 2021;21:29.
46. Barbieri S, Mehta S, Wu B, Bharat C, Poppe K, Jorm L, et al. Predicting cardiovascular risk from national administrative databases using a combined survival analysis and deep learning approach. *Int J Epidemiol*. 2021;dyab258.
47. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. *Nat Rev Cardiol*. 2016;13:350–9.
48. Silverio A, Cavallo P, De Rosa R, Galasso G. Big health data and cardiovascular diseases: a challenge for research, an opportunity for clinical care. *Front Med*. 2019;6:36.
49. Kalantar-Zadeh K, Block G, Horwich T, Fonarow GC. Reverse epidemiology of conventional cardiovascular risk factors in patients with chronic heart failure. *J Am Coll Cardiol*. 2004;43:1439–44.
50. Slagle CE, Conlon FL. Emerging field of cardiomics: high-throughput investigations into transcriptional regulation of cardiovascular development and disease. *Trends Genet*. 2016;32:707–16.
51. Lau E, Wu JC. Omics, big data, and precision medicine in cardiovascular sciences. *Circ Res*. 2018;122:1165–8.
52. V A, Nayar PG, Murugesan R, Mary B, P D, Ahmed SS. Cardiogenbase: a literature-based multi-omics database for major cardiovascular diseases. *PLoS ONE*. 2015;10:e0143188.
53. Vakili D, Radenkovic D, Chawla S, Bhatt DL. Panomics: new databases for advancing cardiology. *Front Cardiovasc Med*. 2021;8:587768.
54. Sarajlic A, Janjic V, Stojkovic N, Radak D, Pržulj N. Network topology reveals key cardiovascular disease genes. *PLoS ONE*. 2013;8:e71537.
55. Zerhouni EA. Clinical research at a crossroads: the NIH roadmap. *J Investig Med*. 2006;54:171–3.
56. Collins FS, Wilder EL, Zerhouni E. Funding transdisciplinary research. *NIH roadmap/common fund at 10 years*. *Science*. 2014;345:274–6.

57. Wang H, Song H, Yang Y, Cao Q, Hu Y, Chen J, et al. Three-dimensional printing for cardiovascular diseases: from anatomical modeling to dynamic functionality. *Biomed Eng Online*. 2020;19:76.
58. Sammani A, Baas AF, Asselbergs FW, Te Riele ASJM. Diagnosis and risk prediction of dilated cardiomyopathy in the era of big data and genomics. *J Clin Med*. 2021;10:921.
59. Lanzer JD, Leuschner F, Kramann R, Levinson RT, Saez-Rodriguez J. Big data approaches in heart failure research. *Curr Heart Fail Rep*. 2020;17:213–24.
60. Chen S, Qi J, Fan S, Qiao Z, Yeo JC, Lim CT. Flexible wearable sensors for cardiovascular health monitoring. *Adv Healthc Mater*. 2021;10:e2100116.
61. Direito A, Rawstorn J, Mair J, Daryabeygi-Khotbehsara R, Maddison R, Tai ES. Multifactorial e- and mHealth interventions for cardiovascular disease primary prevention: protocol for a systematic review and meta-analysis of randomised controlled trials. *Digit Health*. 2019;5:2055207619890480.
62. Gandhi S, Chen S, Hong L, Sun K, Gong E, Li C, et al. Effect of mobile health interventions on the secondary prevention of cardiovascular disease: systematic review and meta-analysis. *Can J Cardiol*. 2017;33:219–31.
63. Wali S, Hussain-Shamsy N, Ross H, Cafazzo J. Investigating the use of mobile health interventions in vulnerable populations for cardiovascular disease management: scoping review. *JMIR Mhealth Uhealth*. 2019;7:e14275.
64. Davis AJ, Parker HM, Gallagher R. Gamified applications for secondary prevention in patients with high cardiovascular disease risk: a systematic review of effectiveness and acceptability. *J Clin Nurs*. 2021;30:3001–10.
65. Prasitlunkum N, Cheungpasitporn W, Chokesuwattanaskul A, Thangjui S, Thongprayoon C, Bathini T, et al. Diagnostic accuracy of smart gadgets/wearable devices in detecting atrial fibrillation: a systematic review and meta-analysis. *Arch Cardiovasc Dis*. 2021;114:4–16.
66. Seetharam K, Shrestha S, Sengupta PP. Cardiovascular imaging and intervention through the lens of artificial intelligence. *Interv Cardiol*. 2021;16:e31.
67. Gonzales RA, Zhang Q, Papiez BW, Werys K, Lukaschuk E, Popescu IA, et al. MOCOnet: robust motion correction of cardiovascular magnetic

resonance T1 mapping using convolutional neural networks. *Front Cardiovasc Med.* 2021;8:768245.

68. Infante T, Cavaliere C, Punzo B, Grimaldi V, Salvatore M, Napoli C. Radiogenomics and artificial intelligence approaches applied to cardiac computed tomography angiography and cardiac magnetic resonance for precision medicine in coronary heart disease: a systematic review. *Circ Cardiovasc Imaging.* 2021;14:1133–46.

69. Antonopoulos AS, Boutsikou M, Simantiris S, Angelopoulos A, Lazaros G, Panagiotopoulos I, et al. Machine learning of native T1 mapping radiomics for classification of hypertrophic cardiomyopathy phenotypes. *Sci Rep.* 2021;11:23596.

70. Rauseo E, Izquierdo Morcillo C, Raisi-Estabragh Z, Gkontra P, Aung N, Lekadir K, et al. New imaging signatures of cardiac alterations in ischaemic heart disease and cerebrovascular disease using CMR radiomics. *Front Cardiovasc Med.* 2021;8:716577.

71. Eysenbach G. Infodemiology: the epidemiology of (mis)information. *Am J Med.* 2002;113:763–5.

72. Eysenbach G. Websites on screening for breast cancer: “infodemiology” studies have surely had their day. *BMJ.* 2004;328:769.

73. Kling JM, Miller VM, Mankad R, Wilansky S, Wu Q, Zais TG, et al. Go Red for Women cardiovascular health-screening evaluation: the dichotomy between awareness and perception of cardiovascular risk in the community. *J Womens Health.* 2013;22:210–8.

74. Suero-Abreu GA, Barajas-Ochoa A, Perez-Peralta A, Rojas E, Berkowitz R. Assessment of the effect of the Go Red for Women campaign on search engine queries for cardiovascular disease in women. *Cardiol Res.* 2020;11:348–52.

75. Dzaye O, Adelhoefer S, Boakye E, Blaha MJ. Cardiovascular-related health behaviors and lifestyle during the COVID-19 pandemic: an infodemiology study. *Am J Prev Cardiol.* 2021;5:100148.

76. Dzaye O, Duebgen M, Berning P, Graham G, Martin SS, Blaha MJ. Understanding myocardial infarction trends during the early COVID-19 pandemic: an infodemiology study. *Intern Med J.* 2021;51:1328–31.

77. Azer SA, AlSwaidan NM, Alshwairikh LA, AlShammari JM. Accuracy and readability of cardiovascular entries on Wikipedia: are they reliable learning resources for medical students? *BMJ Open*. 2015;5:e008187.
78. Jones J, Cassie S, Thompson M, Atherton I, Leslie SJ. Delivering healthcare information via the internet: cardiac patients' access, usage, perceptions of usefulness, and web site content preferences. *Telemed J E Health*. 2014;20:223–7.
79. Okop KJ, Murphy K, Lambert EV, Kedir K, Getachew H, Howe R, et al. Community-driven citizen science approach to explore cardiovascular disease risk perception, and develop prevention advocacy strategies in sub-Saharan Africa: a programme protocol. *Res Involv Engagem*. 2021;7:11.
80. Mensah GA, Cooper RS, Siega-Riz AM, Cooper LA, Smith JD, Brown CH, et al. Reducing cardiovascular disparities through community-engaged implementation research: a National Heart, Lung, and Blood Institute workshop report. *Circ Res*. 2018;122:213–30.
81. Anoushiravani AA, Patton J, Sayeed Z, El-Othmani MM, Saleh KJ. Big data, big research: implementing population health-based research models and integrating care to reduce cost and improve outcomes. *Orthop Clin North Am*. 2016;47:717–24.
82. Anker SD, Agewall S, Borggrefe M, Calvert M, Jaime Caro J, Cowie MR, et al. The importance of patient-reported outcomes: a call for their comprehensive integration in cardiovascular clinical trials. *Eur Heart J*. 2014;35:2001–9.
83. Chen Y, Nagendran M, Gomes M, Wharton PV, Raine R, Lambiase PD. Gaps in patient-reported outcome measures in randomised clinical trials of cardiac catheter ablation: a systematic review. *Eur Heart J Qual Care Clin Outcomes*. 2020;6:234–42.
84. Burns DJ, Arora J, Okunade O, Beltrame JF, Bernardez-Pereira S, Crespo-Leiro MG, et al. International consortium for health outcomes measurement (ICHOM): standardized patient-centered outcomes measurement set for heart failure patients. *JACC Heart Fail*. 2020;8:212–22.
85. Wu PY, Cheng CW, Kaddi CD, Venugopalan J, Hoffman R, Wang MD. -Omic and electronic health record big data analytics for precision medicine. *IEEE Trans Biomed Eng*. 2017;64(2):263–273.
86. Rumsfeld JS, Joynt KE, Maddox TM. Big data analytics to improve cardiovascular care: promise and challenges. *Nat Rev Cardiol*. 2016;13(6):350–359.

87. Ouyang L, Apley DW, Mehrotra S. A design of experiments approach to validation sampling for logistic regression modeling with errorprone medical records. *J Am Med Inform Assoc.* 2016;23(e1):e71–e78.
88. Hoekstra AG, Chopard B, Coster D, Portegies Zwart S, Coveney PV. Multiscale computing for science and engineering in the era of exascale performance. *Philos Trans A Math Phys Eng Sci.* 2019;377(2142):20180144.
89. Cirillo D, Valencia A. Big data analytics for personalized medicine. *Curr Opin Biotechnol.* 2019;58:161–167.
90. Fogel AL, Kvedar JC. Artificial intelligence powers digital medicine. *NPJ Digit Med.* 2018;1:5.
91. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. *Nat Med.* 2019;25(1):44–56.
92. Sardar P, Abbott JD, Kundu A, Aronow HD, Granada JF, Giri J. Impact of artificial intelligence on interventional cardiology: from decision-making aid to advanced interventional procedure assistance. *JACC Cardiovasc Interv.* 2019;12(14):1293–1303.
93. Ong MS, et al. Learning a comorbidity-driven taxonomy of pediatric pulmonary hypertension. *Circ Res.* 2017;121(4):341–353.
94. Lee JH, et al. Phenotypic and genetic heterogeneity among subjects with mild airflow obstruction in COPDGene. *Respir Med.* 2014;108(10):1469–1480.
95. Maron BA, Cockrill BA, Waxman AB, Systrom DM. The invasive cardiopulmonary exercise test. *Circulation.* 2013;127(10):1157–1164.
96. Oldham WM, et al. Network analysis to risk stratify patients with exercise intolerance. *Circ Res.* 2018;122(6):864–876.
97. Antman E, Weiss S, Loscalzo J. Systems pharmacology, pharmacogenetics, and clinical trial design in network medicine. *Wiley Interdiscip Rev Syst Biol Med.* 2012;4(4):367–383.