

Revelar o potencial oculto dos megadados: Capacitar os enfermeiros através da ciência de dados

Unveiling the Hidden Potential of Big Data: Empowering Nurses through Data Science

Vinciya Pandian¹

 <https://orcid.org/0000-0002-2260-1080>

Michael J. Brenner^{2,3}

 <https://orcid.org/0000-0003-4926-0957>

¹ Johns Hopkins School of Nursing,
Baltimore, Estados Unidos da América

² University of Michigan School
of Medicine, Departamento de
Otorrinolaringologia – Cirurgia de
Cabeça e Pescoço, Michigan, Estados
Unidos da América

³ Global Tracheostomy Collaborative,
Illinois, Chicago, Estados Unidos da
América

Introdução

Ao longo do último milénio, o panorama dos cuidados de saúde sofreu uma profunda metamorfose, passando da confiança na intuição e nos remédios tradicionais para uma prática baseada em evidência que tira partido da ciência de dados. Os séculos XIX e XX testemunharam avanços significativos na medicina baseada em evidência, com o advento da investigação científica e a recolha sistemática de dados, bem como a integração da tecnologia na era digital, que deu início a uma nova era de gestão de dados, com os registos de saúde eletrónicos a permitirem a uniformização e a agregação de dados sobre os doentes em grande escala. A sinergia entre os megadados e a ciência de dados está a revolucionar os cuidados de saúde, e este Renascimento oferece uma perspetiva sem precedentes sobre os cuidados prestados aos doentes, os resultados dos tratamentos e a saúde da população. A enfermagem está preparada para capitalizar estas oportunidades, utilizando tecnologias e análises sofisticadas para prestar cuidados personalizados e eficazes. A intersecção entre a enfermagem, os megadados e a ciência de dados assinala uma mudança de paradigma no sentido de um sistema de saúde mais interligado e orientado para os dados, capaz de melhorar a qualidade dos cuidados. Neste editorial, apresentamos a história, o estado atual e o futuro da ciência de dados, destacando a forma como a profissão de enfermagem pode liderar a utilização da inovação digital para catalisar o progresso na educação, na saúde da população, na descoberta de medicamentos e na equidade na saúde. Analisamos ainda as dimensões éticas da privacidade, da integridade dos dados e da discriminação. Enquanto pioneiros na prática baseada em evidência, na colaboração interdisciplinar e na inovação educativa, os enfermeiros estão destinados a moldar o futuro dos cuidados de saúde digitais.

Perspetivas históricas da documentação clínica

As origens da documentação clínica remontam a civilizações antigas. Na antiga Mesopotâmia, as tábuas de argila inscritas com escrita cuneiforme fornecem evidência de alguns dos mais antigos registos médicos conhecidos (Lorkowski & Pokorski, 2022). No Egipto, as informações médicas eram transcritas em rolos de papiro, enquanto na Grécia e em Roma era utilizado pergaminho para documentar as práticas médicas (Lorkowski & Pokorski, 2022). A transição destas práticas antigas para métodos mais sofisticados ocorreu gradualmente ao longo dos séculos. Com o advento da imprensa escrita, o conhecimento médico tornou-se mais amplamente disseminado, permitindo uma documentação uniformizada. No entanto, o aparecimento dos sistemas de registos de saúde eletrónicos transformou a estrutura não só da documentação dos cuidados de saúde, mas também de todos os aspetos da comunicação e da prestação de cuidados de saúde. Estes sistemas permitiram que os profissionais de saúde armazenassem, gerissem e recuperassem informações sobre os doentes e forneceram a arquitetura digital e o repositório necessários para a ciência de dados (Dicuonzo et al., 2022).

Autor de correspondência

Vinciya Pandian

E- mail: vpandia1@jhu.edu



Os megadados referem-se a conjuntos de dados extensos que ultrapassam as capacidades das bases de dados convencionais, sendo que a sua utilização pode transformar os cuidados de saúde (Batko & Ślęzak, 2022). A análise computacional pode revelar padrões, tendências e associações nos cuidados aos doentes, sinalizando aspetos nos dados que só são detetados através de dados populacionais. Esta evolução significa uma abordagem mais ampla e abrangente para extrair informações relevantes da documentação clínica. A ciência de dados integra técnicas analíticas avançadas, algoritmos de aprendizagem automática e modelos estatísticos de forma a obter informações práticas a partir de conjuntos de dados complexos. A mudança na terminologia de *megadados* para *ciência de dados* reflete o reconhecimento de que não é apenas o volume de dados que importa, mas também as metodologias e ferramentas analíticas que lhe são aplicadas. A ciência de dados engloba a análise preditiva, a medicina personalizada e a gestão da saúde da população (Subrahmanya et al., 2022). A ciência de dados está a moldar o futuro dos cuidados de saúde ao revelar novas perspetivas que podem melhorar os resultados, simplificar os processos e melhorar a prestação de cuidados de saúde.

Características dos megadados

Os megadados têm várias dimensões que podem ser resumidas pelos 5 Vs: *volume*, *velocidade*, *variedade*, *veracidade* e *valor*. Cada uma das dimensões pode ser aplicada à prática de enfermagem. O *volume* na ciência de dados envolve conjuntos de dados extensos, desde registos de doentes a ficheiros de imagens médicas, frequentemente medidos em terabytes (Batko & Ślęzak, 2022). A *velocidade* reflete a taxa acumulada de dados dos doentes, necessitando frequentemente de processamento em tempo real para informar as decisões clínicas (Batko & Ślęzak, 2022). A *variedade* abrange o formato e a fonte dos dados, incluindo registos estruturados dos doentes, dados semiestruturados de dispositivos vestíveis e dados não estruturados de narrativas clínicas (Batko & Ślęzak, 2022). A *veracidade* está relacionada com a exatidão e a fiabilidade da informação sobre os doentes, que é fundamental para manter a qualidade dos cuidados (Batko & Ślęzak, 2022). Por último, o *valor* sublinha o objetivo principal de extrair informações valiosas dos dados de enfermagem, facilitando a melhoria contínua dos protocolos e da prestação de cuidados (Batko & Ślęzak, 2022).

A importância da ciência de dados nos cuidados de saúde

A ciência de dados é relevante em diversas áreas da enfermagem, desde o apoio à decisão clínica até aos algoritmos preditivos para antecipar e gerir surtos de doenças. A ciência de dados pode acelerar a descoberta de medicamentos, identificar ineficiências operacionais e apoiar medidas preventivas para beneficiar a saúde da população (Yogesh & Karthikeyan, 2022). Apesar da promessa desta tecnologia, persistem desafios no aproveitamento dos megadados para investigação e disseminação de resultados de qualidade. Os enfermeiros devem documentar e recolher dados de forma consistente para apoiar a investigação, promovendo a colaboração e a tradução do conhecimento. A utilização pioneira da estatística por Florence Nightingale durante a Guerra da Crimeia destaca o poder transformador dos dados (Bradshaw, 2020). Florence Nightingale demonstrou que a tomada de decisões deve ser feita com base na evidência e não na perceção ou na tradição. Atualmente, os enfermeiros têm a oportunidade de utilizar a ciência de dados para melhorar os resultados no domínio da saúde (Bradshaw, 2020). A ciência de dados está a transformar os cuidados de saúde, nomeadamente através de dispositivos portáteis e tecnologias vestíveis (Uddin & Syed-Abdul, 2020). Os modelos individualizados baseados em trajetórias pessoais de saúde permitem cuidados personalizados. Num cenário de rápido fluxo de dados provenientes de bases de dados, dispositivos vestíveis e outras fontes, os enfermeiros desempenham um papel fundamental nas análises significativas, tanto a nível individual como populacional.

Análise de dados em enfermagem

De acordo com o *National Institute of Standards and Technology*, a análise de dados envolve o processamento e a manipulação sistemáticos de dados para revelar padrões, relações, tendências históricas e previsões de comportamentos e eventos futuros (National Institute of Standards and Technology, 2015). Portanto, a análise de dados é muito mais ampla do que as análises estatísticas, englobando a preparação da informação, a análise de padrões, a síntese de conhecimentos e a criação de valor (National Institute of Standards and Technology, 2015). O processo analítico começa com a recolha, extração, processamento e transformação de dados e prossegue para a análise e a criação de modelos e algoritmos, culminando na interpretação e elaboração de relatórios que podem informar a tomada de decisão (National Institute of Standards and Technology, 2015). A análise engloba os métodos, a implementação e os resultados necessários para sintetizar a informação e orientar a tomada de decisão.

Tipos de análises

A análise pode ser descritiva, diagnóstica, preditiva ou prescritiva. A *análise descritiva* inclui a utilização de inteligência empresarial ou extração de dados para explicar ocorrências passadas e fornecer informações sobre eventos históricos sem aprofundar as causas subjacentes (Sarker, 2021). A *análise diagnóstica* procura descobrir as razões subjacentes aos eventos, comparando dados históricos com outras fontes e fornecendo informações aprofundadas sobre problemas específicos (Sarker, 2021). A *análise preditiva* utiliza modelos estatísticos e projeções para prever tendências futuras, muitas vezes com base em algoritmos de aprendizagem automática ou de aprendizagem profunda (*Deep Learning*; Sarker, 2021). A *análise prescritiva* utiliza a otimização e a simulação para recomendar ações com o objetivo de

eliminar problemas futuros ou capitalizar tendências promissoras através de ferramentas e tecnologias sofisticadas e uma análise cuidadosa dos dados históricos e externos (Sarker, 2021).

Fontes existentes para a ciência de dados em enfermagem

No domínio da ciência de dados em enfermagem, várias fontes revelam-se indispensáveis para os enfermeiros que necessitam de informações valiosas. A MarketScan, uma base de dados longitudinal, fornece informação sobre pedidos de reembolso de cuidados de saúde a nível individual, facilitando estudos aprofundados sobre cuidados cirúrgicos e gestão de doenças (Hirsch & Pitak-Arnopp, 2023; Premo et al., 2023; Wadhwa et al., 2023; Xie et al., 2023). O *Healthcare Cost and Utilization Project* (HCUP) inclui bases de dados estaduais e nacionais, oferecendo informações abrangentes sobre altas hospitalares, idas ao serviço de urgência e consultas de ambulatório em centros de cirurgia (Ma et al., 2023; Matabele et al., 2023). As bases de dados nacionais, tais como a *Kids' Inpatient Database*, a *National Inpatient Sample* e as *National Readmission Databases*, fornecem dados sobre internamento de base familiar, contribuindo para as estimativas nacionais sobre a carga de doença (Metcalfe et al., 2019). Os registos especializados, tais como o registo de dados clínicos *Reg-ent* da *American Academy of Otolaryngology-Head and Neck Surgery*, fornecem dados clínicos específicos sobre especialidades, permitindo avaliações de resultados (Saraswathula et al., 2023; Schmalbach et al., 2021). Além disso, iniciativas como o *Global Tracheostomy Collaborative Registry*, que recolhe dados de mais de 50 hospitais a nível mundial, utilizam a ciência de dados para melhorar a compreensão e o tratamento dos doentes com traqueostomia (Brenner et al., 2020). Estas diferentes fontes ajudam os enfermeiros coletivamente nos seus esforços ao nível da ciência de dados, fornecendo uma base rica para análise e inovação nos cuidados de saúde.

Desafios da ciência de dados nos cuidados de saúde

As considerações éticas são uma pedra angular do aproveitamento dos megadados e da ciência de dados para os avanços nos cuidados de saúde. As considerações específicas dizem respeito à privacidade, segurança e equidade dos dados. As partes interessadas - dos doentes aos profissionais de saúde - devem adotar uma perspetiva de mente aberta que inclua a vontade de analisar criticamente e aprender com a ciência de dados. Garantir a exatidão e a consistência dos dados é fundamental, sendo que é necessária uma otimização contínua para evitar perceções e decisões erradas (Osama et al., 2023). Além disso, trabalhar com diversos conjuntos de dados implica navegar em silos de dados, questões de interoperabilidade e redução da discriminação (Osama et al., 2023). A integração de diferentes sistemas de armazenamento de dados requer mapeamento, transformação, consolidação e garantia de compatibilidade para uma análise e uma tomada de decisão eficazes (Osama et al., 2023). Esses esforços são fundamentais para uma análise representativa que pode ajudar a reduzir a discriminação e a perpetuar as desigualdades na saúde.

Avanços tecnológicos e tendências emergentes na ciência de dados

A ciência de dados capacite os enfermeiros para prestarem cuidados personalizados e defenderem a melhoria das práticas de cuidados. A análise preditiva permite a identificação proactiva de problemas de saúde iminentes e a expansão da telessaúde e da monitorização remota pode melhorar a monitorização de doenças crónicas. O crescimento da tecnologia vestível acrescenta outra dimensão, permitindo que os enfermeiros utilizem dados de dispositivos e sensores para monitorizar os sinais vitais, os níveis de atividade e a adesão ao plano de tratamento (Uddin & Syed-Abdul, 2020). A medicina genómica é outra fronteira para a ciência de dados, permitindo aos enfermeiros desenvolver e implementar planos de tratamento personalizados (Hassan et al., 2022). A ciência de dados é também fundamental na gestão da saúde da população, permitindo aos enfermeiros identificar populações em risco e criar intervenções direcionadas para melhorar a saúde das comunidades. O Processamento de Linguagem Natural pode extrair dados de notas clínicas e narrativas não estruturadas, facilitando uma compreensão mais aprofundada (Aramaki et al., 2022). À medida que o sector avança, os enfermeiros devem reconhecer as promessas e os perigos dos conjuntos de dados e das capacidades analíticas em constante expansão. A cibersegurança é fundamental para garantir a privacidade dos dados dos doentes, sendo que é necessária uma vigilância constante para salvaguardar as informações sensíveis no sector da saúde (Cremer et al., 2022). O advento da medicina de precisão, os testes genómicos e a integração da Inteligência Artificial na análise de dados estão a transformar os cuidados de enfermagem (Johnson et al., 2021). Os ensaios clínicos são cada vez mais realizados virtualmente. Os *smartphones* permitem o acesso contínuo aos ensaios e as crescentes capacidades de computação dos dispositivos abrem novas possibilidades. A aprendizagem automática é um importante catalisador na evolução da ciência de dados, lançando as bases para uma maior integração da Inteligência Artificial nos cuidados de saúde. Os enfermeiros serão fundamentais na aceitação da mudança e na adaptação a novas dinâmicas necessárias para navegar num cenário em constante mudança ao nível dos cuidados de saúde.

Dados e o futuro da enfermagem

Ao aceitarem a integração dos megadados nos cuidados de saúde, os enfermeiros podem liderar descobertas inovadoras e desenvolver ferramentas baseadas em dados que revolucionem os cuidados prestados aos doentes. A ciência de dados pode servir de base para as tecnologias imersivas necessárias para acelerar a aprendizagem e melhorar o desempenho. A ciência de dados ajuda os futuros profissionais de saúde a navegarem com confiança e competência

nas complexidades de um panorama de cuidados de saúde rico em dados. Os enfermeiros são fundamentais nos debates políticos e podem defender a utilização ética dos dados, a privacidade dos doentes e a utilização responsável dos megadados. A colaboração interdisciplinar com profissionais de diversas áreas promove a inovação e cria soluções que ultrapassam as fronteiras tradicionais. Através da aprendizagem ao longo da vida e da atualização constante sobre os avanços na ciência de dados, os enfermeiros estarão preparados para enfrentar os desafios do futuro.

Conclusão

O papel transformador dos megadados e da ciência de dados nos cuidados de saúde posicionou os enfermeiros como líderes na definição do futuro dos cuidados de saúde. Ao alavancar os avanços tecnológicos no contexto da educação, da colaboração interdisciplinar e da utilização ética dos dados, os enfermeiros irão assumir a liderança na investigação, na inovação e na *advocacy*. Para tal, os enfermeiros devem desenvolver competências de acesso aos dados e interpretação das análises de dados, bem como adotar tendências emergentes. Os enfermeiros estão numa posição em que podem revolucionar os cuidados prestados aos doentes através da aplicação judiciosa dos megadados e da ciência de dados.

Referências bibliográficas

- Aramaki, E., Wakamiya, S., Yada, S., & Nakamura, Y. (2022). Natural Language Processing: from Bedside to Everywhere. *Yearb Med Inform*, 31(1), 243-253. <https://doi.org/10.1055/s-0042-1742510>
- Batko, K., & Ślęzak, A. (2022). The use of Big Data Analytics in healthcare. *J Big Data*, 9(1), 3. <https://doi.org/10.1186/s40537-021-00553-4>
- Bradshaw, N. A. (2020). Florence Nightingale (1820-1910): An Unexpected Master of Data. *Patterns (N Y)*, 1(2), 100036. <https://doi.org/10.1016/j.patter.2020.100036>
- Brenner, M. J., Pandian, V., Milliren, C. E., Graham, D. A., Zaga, C., Morris, L. L., Bedwell, J. R., Das, P., Zhu, H., Lee, Y. A. J., Peltz, A., Chin, K., Schiff, B. A., Randall, D. M., Swords, C., French, D., Ward, E., Sweeney, J. M., Warrillow, S. J., . . . Roberson, D. W. (2020). Global Tracheostomy Collaborative: data-driven improvements in patient safety through multidisciplinary teamwork, standardisation, education, and patient partnership. *Br J Anaesth*, 125(1), e104-e118. <https://doi.org/10.1016/j.bja.2020.04.054>
- Cremer, F., Sheehan, B., Fortmann, M., Kia, A. N., Mullins, M., Murphy, F., & Materne, S. (2022). Cyber risk and cybersecurity: a systematic review of data availability. *Geneva Pap Risk Insur Issues Pract*, 47(3), 698-736. <https://doi.org/10.1057/s41288-022-00266-6>
- Dicuonzo, G., Galeone, G., Shini, M., & Massari, A. (2022). Towards the Use of Big Data in Healthcare: A Literature Review. *Healthcare (Basel)*, 10(7). <https://doi.org/10.3390/healthcare10071232>
- Hassan, M., Awan, F. M., Naz, A., deAndrés-Galiana, E. J., Alvarez, O., Cernea, A., Fernández-Brillet, L., Fernández-Martínez, J. L., & Kloczkowski, A. (2022). Innovations in Genomics and Big Data Analytics for Personalized Medicine and Health Care: A Review. *Int J Mol Sci*, 23(9). <https://doi.org/10.3390/ijms23094645>
- Hirsch, J., & Pitak-Arnop, P. (2023). [Epidemiology and treatment of traumatic facial nerve palsy following skull base fractures: Results from the IBM MarketScan commercial database (2006-2019)]. *Unfallchirurgie (Heidelberg)*. <https://doi.org/10.1007/s00113-023-01367-0>
- Johnson, K. B., Wei, W. Q., Weeraratne, D., Frisse, M. E., Misulis, K., Rhee, K., Zhao, J., & Snowdon, J. L. (2021). Precision Medicine, AI, and the Future of Personalized Health Care. *Clin Transl Sci*, 14(1), 86-93. <https://doi.org/10.1111/cts.12884>
- Lorkowski, J., & Pokorski, M. (2022). Medical Records: A Historical Narrative. *Biomedicines*, 10(10). <https://doi.org/10.3390/biomedicines10102594>
- Ma, J., Johnson, E. A., & McCrory, B. (2023). Predicting risk factors for pediatric mortality in clinical trial research: A retrospective, cross-sectional study using a Healthcare Cost and Utilization Project database. *J Clin Transl Sci*, 7(1), e211. <https://doi.org/10.1017/cts.2023.634>
- Matabele, M. N., Cheng, C., Venkatesh, M., Barr, S., Ellefson, J., Beninati, M., Lobeck, I. N., & Puricelli, M. D. (2023). Perinatal airway management in neonatal goiter: A healthcare cost and utilization project (HCUP) kids' inpatient database analysis. *Int J Pediatr Otorhinolaryngol*, 175, 111767. <https://doi.org/10.1016/j.ijporl.2023.111767>
- Metcalfe, D., Zogg, C. K., Haut, E. R., Pawlik, T. M., Haider, A. H., & Perry, D. C. (2019). Data resource profile: State Inpatient Databases. *Int J Epidemiol*, 48(6), 1742-1742h. <https://doi.org/10.1093/ije/dyz117>
- National Institute of Standards and Technology. (2015). *NIST Big Data Interoperability Framework: Volume 1, Definitions* National Institute of Standards and Technology. <https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-1.pdf>
- Osama, M., Ateya, A. A., Sayed, M. S., Hammad, M., Pławiak, P., Abd El-Latif, A. A., & Elsayed, R. A. (2023). Internet of Medical Things and Healthcare 4.0: Trends, Requirements, Challenges, and Research Directions. *Sensors (Basel)*, 23(17). <https://doi.org/10.3390/s23177435>
- Premo, H., Gordee, A., Lee, H. J., Scales, C. D., Moul, J. W., & Peterson, A. (2023). Disparities in Prostate Cancer Screening for Transgender Women: An Analysis of the MarketScan Database. *Urology*, 176, 237-242. <https://doi.org/10.1016/j.urology.2023.03.016>
- Saraswathula, A., Roy, S., Blythe, W. R., Gourin, C. G., & Boss, E. F. (2023). The Unrealized Potential of the Reg-ent ENT Clinical Data Registry. *JAMA Otolaryngol Head Neck Surg*, 149(8), 659-661. <https://doi.org/10.1001/jamaoto.2023.1389>
- Sarker, I. H. (2021). Data Science and Analytics: An Overview from Data-Driven Smart Computing, Decision-Making and Applications Perspective. *SN Comput Sci*, 2(5), 377. <https://doi.org/10.1007/s42979-021-00765-8>
- Schmalbach, C. E., Brereton, J., Bowman, C., & Denny, J. C., 3rd. (2021). American Academy of Otolaryngology-Head and Neck



- Surgery/Foundation Registry: Purpose, Properties, and Priorities. *Otolaryngol Head Neck Surg*, 164(5), 964-971. <https://doi.org/10.1177/0194599820984135>
- Subrahmanya, S. V. G., Shetty, D. K., Patil, V., Hameed, B. M. Z., Paul, R., Smriti, K., Naik, N., & Somani, B. K. (2022). The role of data science in healthcare advancements: applications, benefits, and future prospects. *Ir J Med Sci*, 191(4), 1473-1483. <https://doi.org/10.1007/s11845-021-02730-z>
- Uddin, M., & Syed-Abdul, S. (2020). Data Analytics and Applications of the Wearable Sensors in Healthcare: An Overview. *Sensors (Basel)*, 20(5). <https://doi.org/10.3390/s20051379>
- Wadhwa, H., Varshneya, K., Stienen, M. N., & Veeravagu, A. (2023). Do Epidural Steroid Injections Affect Outcomes and Costs in Cervical Degenerative Disease? A Retrospective MarketScan Database Analysis. *Global Spine J*, 13(7), 1812-1820. <https://doi.org/10.1177/21925682211050320>
- Xie, F., Beukelman, T., Sun, D., Yun, H., & Curtis, J. R. (2023). Identifying inpatient mortality in MarketScan claims data using machine learning. *Pharmacoepidemiol Drug Saf*, 32(11), 1299-1305. <https://doi.org/10.1002/pds.5658>
- Yogesh, M. J., & Karthikeyan, J. (2022). Health Informatics: Engaging Modern Healthcare Units: A Brief Overview. *Front Public Health*, 10, 854688. <https://doi.org/10.3389/fpubh.2022.854688>