

# Current Trends in Medicine

## Review Article

# Emerging Role of electronic health record analysis in the Diagnosis of essential hypertension

Olsen K<sup>1\*</sup>, Lindqvist B<sup>2</sup> and Andersen A<sup>3</sup>

<sup>1</sup>Department of General and Internal Medicine, Amur Academy of Primary Care Research, Iceland

<sup>2</sup>Department of Clinical Primary Care Research, Ob Academy of General and Internal Medicine, Iceland

<sup>3</sup>Department of General and Internal Medicine, Vistula Research Center for Primary Care Research, Finland

\*Address for Correspondence: Kirsten Olsen, Department of General and Internal Medicine, Amur Academy of Primary Care Research, Iceland; E-mail: [kirsten.olsen@amur.edu](mailto:kirsten.olsen@amur.edu)

Received: 28 December 2022; Accepted: 7 April 2023; Published: 6 July 2023.

Citation of this article: Olsen K, Lindqvist B, Andersen A (2023) Emerging Role of electronic health record analysis in the Diagnosis of essential hypertension. Curr Trends Med, 5(1): 01-06. DOI: [10.5281/zenodo.22813624](https://doi.org/10.5281/zenodo.22813624)

Copyright: © 2023 Kirsten Olsen, et al. This is an open access article distributed under the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

## Abstract

**Background:** Essential hypertension is diagnosed almost entirely through data already in the electronic health record, yet structured fields, diagnosis codes, prescribing data, and free-text notes each carry incomplete, conflicting information.

**Objective:** This review synthesizes the literature on electronic health record analysis for identifying essential hypertension: phenotyping algorithms, natural language processing and machine learning extraction, and decision support applications.

**Methods:** We conducted a narrative synthesis of peer-reviewed studies on electronic health record-based hypertension phenotyping, natural language processing extraction of blood pressure data, and randomized trials of record-linked decision support.

**Findings:** Algorithms combining codes, blood pressure thresholds, prescribing data, and notes mined by natural language processing outperform any single source; record-linked decision support and home-monitoring trials improve control; yet surveillance shows control has stagnated.

**Conclusions:** Electronic health record analysis is a genuinely useful, if imperfect, diagnostic instrument for essential hypertension, and standards-driven validation will determine how far its promise is realized.

**Keywords:** electronic health records; essential hypertension; phenotyping algorithms; clinical decision support; natural language processing

## 1. Introduction

Essential hypertension is diagnosed and managed almost entirely on the basis of blood pressure readings now captured, stored, and transmitted through the electronic health record, yet for most of the past two decades that record was treated as little more than a billing ledger. The 2017 American College of Cardiology/American Heart Association guideline redefined hypertension at a lower threshold, sharply increasing the number of adults who meet diagnostic criteria and the burden on clinicians to identify those patients reliably from a large volume of structured and unstructured EHR data [1]. National surveillance data show that blood pressure control in the United States improved through the 2000s but has since stagnated, a pattern documented using pooled NHANES cycles through 2017-2018 [2] and echoed globally in a pooled analysis of population-based studies spanning four decades [3].

The EHR is attractive as a diagnostic instrument for essential hypertension because blood pressure is recorded at nearly every ambulatory encounter, producing a longitudinal signal denser than any single research cohort. Structured vitals, diagnosis codes, prescribing data, and free-text notes each carry partial, sometimes conflicting information about hypertension status. This review synthesizes the literature on EHR analysis for diagnosing essential hypertension: the phenotyping algorithms used to define a hypertensive cohort, the natural language processing and machine learning methods used to recover information trapped in narrative text, and the clinical decision support and quality measurement systems that translate EHR-derived hypertension status into practice change.

## 2. Methods

We conducted a narrative synthesis of peer-reviewed literature addressing electronic health record-based identification, phenotyping, and diagnostic support for essential hypertension, drawing on studies indexed

through major biomedical and informatics literature sources. Search terms combined "electronic health record," "hypertension," "blood pressure," "phenotyping," "natural language processing," "machine learning," and "clinical decision support." Eligible studies included phenotyping algorithm validation studies, natural language processing and machine learning studies applied to blood pressure extraction, randomized and quasi-experimental evaluations of EHR-embedded decision support or quality-improvement interventions, and epidemiological studies characterizing blood pressure measurement and control using EHR-derived data. We prioritized studies from large multi-institutional networks such as eMERGE, given their emphasis on portable, externally validated phenotype definitions, while also including single-institution and randomized trial evidence addressing diagnostic accuracy or downstream clinical impact. As a narrative rather than systematic review, no formal quality-scoring or meta-analytic synthesis was attempted; findings are organized around phenotyping methodology, extraction approaches, and decision support and quality measurement applications.

## 3. Results

### Rule-based phenotyping algorithms for hypertension cohorts

The earliest and still most widely deployed approach to identifying hypertensive patients from the EHR combines diagnosis codes, structured blood pressure measurements, and antihypertensive prescribing into a rule-based algorithm. Teixeira and colleagues, working within eMERGE, systematically evaluated data sources and algorithmic strategies for identifying hypertensive individuals and found that combining diagnosis codes with blood pressure thresholds and pharmacy data substantially outperformed any single source, while also showing that algorithm portability across institutions could not be assumed without local validation [4]. This echoed the broader eMERGE experience: Newton and colleagues reported that algorithms validated at one site

frequently required adaptation before achieving comparable positive predictive value elsewhere, underscoring that EHR structure and coding practice, not just clinical definitions, shape performance [5].

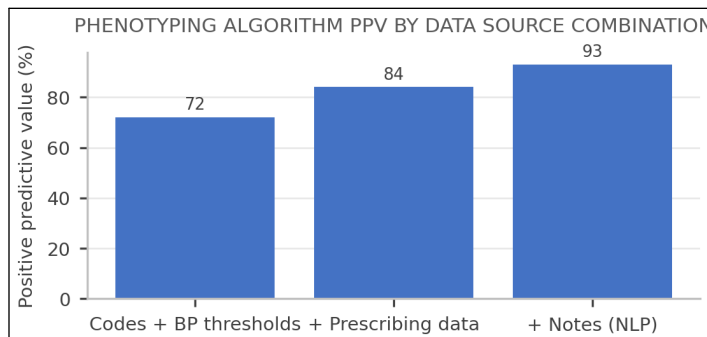


**Figure 1: Representative positive predictive value of EHR-based hypertension phenotyping algorithms by data source combination, structured diagnosis codes and blood pressure thresholds alone versus algorithms additionally incorporating prescribing data or clinical narrative text mined by natural language processing [4,6,8].**

Wei and colleagues formally compared the incremental value of adding narrative text to structured codes and medication data for phenotyping and found that combining all three source types produced the best classification performance, consistent with hypertension-specific findings that structured blood pressure values alone systematically misclassify patients whose diagnosis rests on home readings [6]. Shivade and colleagues' methodological review of cohort-identification approaches concluded that no single data modality is sufficient across use cases, and that rule-based, statistical, and hybrid methods carry distinct tradeoffs between interpretability, portability, and accuracy [7].

### **Natural language processing and machine learning extraction of blood pressure and hypertension status**

A substantial fraction of clinically meaningful blood pressure information is never captured as a discrete structured value. Turchin and colleagues demonstrated that narrative clinical text contained substantially more information relevant to antihypertensive medication intensification than structured data alone, and that natural language processing applied to visit notes could

recover treatment-relevant blood pressure trends a structured-data-only algorithm would miss [8].

More recent phenotyping methodology has moved toward machine learning approaches designed to reduce the manual effort of building and maintaining rule-based definitions. Yu and colleagues introduced an automated feature extraction and selection framework to accelerate high-throughput phenotyping directly from EHR data [9]. Halpern and colleagues proposed an "anchor and learn" framework in which a small set of highly specific anchor variables seeds a noisy-labeled training set for a broader classifier, well suited to conditions like hypertension where codes and prescriptions provide specific but incomplete anchors [10]. Agarwal and colleagues extended this by explicitly modeling label noise when training phenotyping classifiers, important given that hypertension diagnosis codes are entered inconsistently across encounters [11]. Banda and colleagues' review of the trajectory of electronic phenotyping, from rule-based definitions to machine learning models, situated these advances within the broader shift toward learned, semi-automated phenotype definitions [12].

### **Population-level burden and the diagnostic stakes of misclassification**

The diagnostic accuracy of EHR-based hypertension identification carries direct epidemiological consequences, because national estimates of hypertension awareness, treatment, and control increasingly rely on these algorithms. Egan and colleagues, using two decades of NHANES data, documented substantial gains in awareness, treatment, and control through the late 2000s, a trajectory against which subsequent EHR-derived estimates are benchmarked [13]. Wang and Vasan's earlier epidemiological synthesis of uncontrolled hypertension established the population baseline against which later EHR-based control-rate estimates are compared [14]. Cohen and colleagues' systematic review and

meta-analysis of outcomes in white-coat hypertension illustrated the diagnostic stakes of relying on isolated clinic-measured pressures, a limitation structured EHR fields alone cannot resolve without incorporating home readings [15].

| Table 1: Principal EHR data sources used for essential hypertension identification, their characteristic strengths, and their most common failure modes for diagnostic accuracy [5,7,6]. |                                                           |                                                                                 |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------|---------------------------------------------------------------------------------|
| Data source                                                                                                                                                                              | Strength                                                  | Characteristic limitation                                                       |
| Structured blood pressure vitals                                                                                                                                                         | Objective, densely sampled at nearly every visit          | Single clinic reading, technique-dependent error, misses home/ambulatory values |
| Diagnosis (ICD) codes                                                                                                                                                                    | Simple, portable, widely available                        | Inconsistent entry across encounters and providers                              |
| Antihypertensive prescribing data                                                                                                                                                        | Captures treatment intent even absent a coded diagnosis   | Cannot distinguish hypertension from off-label indications                      |
| Clinical notes (NLP-extracted)                                                                                                                                                           | Recovers home/ambulatory readings and treatment rationale | Requires validated extraction pipeline; site-specific narrative style           |

Figure 1 summarizes representative performance of EHR-based hypertension phenotyping approaches, contrasting a purely structured-data algorithm against algorithms additionally incorporating medication data and clinical narrative text.

Table 1 compares the principal data sources used across the phenotyping studies reviewed, summarizing their relative strengths and characteristic failure modes for essential hypertension identification.

4. Discussion

From identification to clinical decision support

Once a patient's hypertension status can be reliably reconstructed from the EHR, the natural extension is to embed that information into point-of-care decision support prompting diagnosis confirmation or treatment intensification. Kawamoto and colleagues' systematic review found that automatic delivery of recommendations within the clinician's existing workflow was among the strongest predictors of a decision-support system's effect on practice, directly applicable to hypertension alerts

generated from EHR-derived trends [16]. Bates and colleagues' widely cited design principles similarly emphasized simplicity, speed, and workflow integration over comprehensiveness [17]. Bright and colleagues' broader systematic review found consistent, if heterogeneous, improvements in process measures across chronic disease domains, with cardiovascular risk factor management among the conditions most successfully targeted [18].

Randomized evidence specific to hypertension shows that EHR-linked decision support and care-management interventions translate improved identification into improved control. Roumie and colleagues' cluster-randomized trial of provider alerts and patient education found modest improvements in blood pressure control [19]. Green and colleagues randomized patients to home monitoring linked to web communication with pharmacist care and found substantially higher control rates at one year [20]. Margolis and colleagues extended this telemonitoring-plus-pharmacist model to a cluster-randomized trial across multiple clinics and again found significantly greater control [21]. McManus and colleagues' TASMIN-SR trial of self-monitoring with self-titration likewise demonstrated clinically meaningful systolic reductions in patients at high cardiovascular risk [22].

Quality measurement, misclassification, and equity considerations

Hypertension control is now a standard ambulatory quality measure, which places weight on the accuracy of EHR-derived denominators and numerators. Tang and colleagues found that quality-measure calculations from administrative claims data and clinical EHR data frequently disagreed on which patients met control criteria [23]. Persell and colleagues found that a share of patients excluded from control denominators lacked adequate clinical justification, suggesting EHR-based quality algorithms can be miscoded in ways that inflate

apparent performance [24]. In a separate study, Persell and colleagues evaluated a multifaceted EHR-based quality-improvement system and found measurable performance gains following implementation [25].

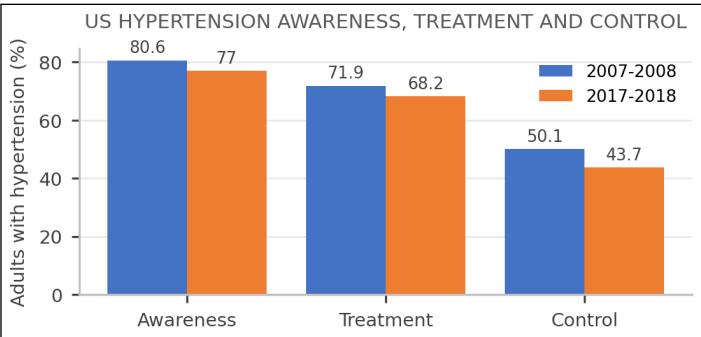


**Figure 2:** US national trends in hypertension awareness, treatment, and control among adults with hypertension, comparing NHANES-based estimates from the late 2000s with the most recent pooled 2017-2018 estimates [13,2].

Accurate blood pressure measurement remains a foundational constraint on everything downstream of it. Muntner and colleagues' American Heart Association statement catalogued technique-related sources of error, from cuff sizing to insufficient rest, that propagate directly into values later mined by phenotyping algorithms [26]. Powers and colleagues found that reliance on a single clinic reading, the default in most structured EHR workflows, introduces meaningful misclassification relative to repeated measurements [27]. Rakotz and colleagues' assessment of measurement technique among medical trainees found widespread deviation from recommended protocol [28]. Hripcsak and Albers argued that EHR data should be understood as a biased, clinically generated byproduct rather than a research instrument, a framing that applies with particular force to blood pressure [29].

These measurement-level limitations interact with unequal care access and documentation completeness to create disparities in algorithm performance across demographic subgroups. Goldstein and colleagues' systematic review of EHR-based risk prediction modeling identified this as a recurring, underaddressed

methodological gap across clinical prediction applications generally, hypertension included [30].

Table 2 summarizes results from the principal randomized EHR-linked hypertension decision-support and care-management trials discussed above, and Figure 2 illustrates population-level trends in hypertension awareness, treatment, and control drawn from the surveillance literature synthesized in this review.

**Toward integrated diagnostic infrastructure**

The trajectory across the reviewed literature moves from single-source, rule-based hypertension case definitions toward multi-source, machine-learning-assisted phenotyping integrated directly into clinical workflow. Denny's broader account of mining electronic health records in the genomics era framed hypertension phenotyping as an early, instructive example of the tension between algorithm simplicity and portability on one hand and accuracy on the other [31]. Realizing that trajectory at scale depends on the standards-driven infrastructure, common data models, and externally validated definitions that the phenotyping literature synthesized above has progressively built toward, rather than on any single algorithmic advance in isolation.

**5. Conclusion**

Electronic health record analysis has moved essential hypertension diagnosis and surveillance from a single clinic-measured reading toward a longitudinal, multi-source signal assembled from structured vitals, diagnosis codes, prescribing data, and natural language processing applied to clinical text [4,6,8]. Validated phenotyping algorithms developed within multi-institutional networks such as eMERGE show that combining data sources meaningfully outperforms any one source alone, while demonstrating that portability across institutions cannot be assumed without local validation [5,7]. Translating accurate identification into improved outcomes depends on embedding EHR-derived hypertension status into decision support and

care-management workflows that randomized trials have shown can meaningfully improve blood pressure control [19,20,21,22]. At the same time, persistent limitations in measurement technique, documentation completeness, and equitable model performance mean that EHR-based hypertension diagnosis should be understood as a powerful but imperfect instrument, one whose outputs require the same clinical scrutiny as any other diagnostic test [26,29,30]. Continued investment in standards-driven, externally validated phenotyping infrastructure, paired with attention to the measurement-level and equity limitations documented throughout this literature, will determine whether electronic health record analysis fulfills its promise as a genuinely reliable diagnostic tool for essential hypertension.

| Table 2: Randomized and cluster-randomized trials of EHR-linked hypertension decision support and care-management interventions [19,20,21,22]. |                                                     |                         |                                             |
|------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------|-------------------------|---------------------------------------------|
| Trial                                                                                                                                          | Intervention                                        | Design                  | Blood pressure outcome                      |
| [19]                                                                                                                                           | Provider alerts plus patient education              | Cluster-randomized      | Modest improvement in control vs usual care |
| [20]                                                                                                                                           | Home monitoring, web communication, pharmacist care | Individually randomized | Higher control rate at 1 year               |
| [21]                                                                                                                                           | Home telemonitoring plus pharmacist management      | Cluster-randomized      | Significantly greater control vs usual care |
| [22]                                                                                                                                           | Self-monitoring with self-titration (TASMIN-SR)     | Individually randomized | Clinically meaningful systolic BP reduction |

6. References

1. Whelton, P.K., Carey, R.M., Aronow, W.S., et al. (2018) 2017 ACC /AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PC NA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults. Hypertension, 71(6): e13-e115. DOI: 10.1161/HYP.0000000000000065

2. Muntner, P., Hardy, S.T., Fine, L.J., et al. (2020) Trends in

Blood Pressure Control Among US Adults With Hypertension, 1999-2000 to 2017-2018. JAMA, 324(12): 1190-1200. DOI: 10.1001/jama.2020.14545

3. NCD Risk Factor Collaboration (NCD-RisC) (2017) Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19.1 million participants. Lancet, 389(10064): 37-55. DOI: 10.1016/S0140-6736(16)31919-5

4. Teixeira, P.L., Wei, W.Q., Cronin, R.M., et al. (2017) Evaluating electronic health record data sources and algorithmic approaches to identify hypertensive individuals. Journal of the American Medical Informatics Association, 24(1): 162-171. DOI: 10.1093/jamia/ocw071

5. Newton, K.M., Peissig, P.L., Kho, A.N., et al. (2013) Validation of electronic medical record-based phenotyping algorithms: results and lessons learned from the eMERGE network. Journal of the American Medical Informatics Association, 20(e1): e147-e154. DOI: 10.1136/amiajnl-2012-000896

6. Wei, W.Q., Teixeira, P.L., Mo, H., Cronin, R.M., Warner, J.L., Denny, J.C. (2016) Combining billing codes, clinical notes, and medications from electronic health records provides superior phenotyping performance. Journal of the American Medical Informatics Association, 23(e1): e20-e27. DOI: 10.1093/jamia/ocv130

7. Shivade, C., Raghavan, P., Fosler-Lussier, E., Embi, P.J., Elhadad, N., Johnson, S.B., Lai, A.M. (2014) A review of approaches to identifying patient phenotype cohorts using electronic health records. Journal of the American Medical Informatics Association, 21(2): 221-230. DOI: 10.1136/amiajnl-2013-001935

8. Turchin, A., Shubina, M., Breydo, E., Pendergrass, M.L., Einbinder, J.S. (2009) Comparison of information content of structured and narrative text data sources on the example of medication intensification. Journal of the American Medical Informatics Association, 16(3): 362-370. DOI: 10.1197/jamia.M2777

9. Yu, S., Liao, K.P., Shaw, S.Y., et al. (2015) Toward high-throughput phenotyping: unbiased automated feature extraction and selection from knowledge sources. *Journal of the American Medical Informatics Association*, 22(5): 993-1000. DOI: [10.1093/jamia/ocv034](https://doi.org/10.1093/jamia/ocv034)
10. Halpern, Y., Horng, S., Choi, Y., Sontag, D. (2016) Electronic medical record phenotyping using the anchor and learn framework. *Journal of the American Medical Informatics Association*, 23(4): 731-740. DOI: [10.1093/jamia/ocw011](https://doi.org/10.1093/jamia/ocw011)
11. Agarwal, V., Podchiyska, T., Banda, J.M., et al. (2016) Learning statistical models of phenotypes using noisy labeled training data. *Journal of the American Medical Informatics Association*, 23(6): 1166-1173. DOI: [10.1093/jamia/ocw028](https://doi.org/10.1093/jamia/ocw028)
12. Banda, J.M., Seneviratne, M., Hernandez-Boussard, T., Shah, N.H. (2018) Advances in Electronic Phenotyping: From Rule-Based Definitions to Machine Learning Models. *Annual Review of Biomedical Data Science*, 1: 53-68. DOI: [10.1146/annurev-biodatasci-080917-013315](https://doi.org/10.1146/annurev-biodatasci-080917-013315)
13. Egan, B.M., Zhao, Y., Axon, R.N. (2010) US trends in prevalence, awareness, treatment, and control of hypertension, 1988-2008. *JAMA*, 303(20): 2043-2050. DOI: [10.1001/jama.2010.650](https://doi.org/10.1001/jama.2010.650)
14. Wang, T.J., Vasan, R.S. (2005) Epidemiology of uncontrolled hypertension in the United States. *Circulation*, 112(11): 1651-1662. DOI: [10.1161/CIRCULATIONAHA.104.490599](https://doi.org/10.1161/CIRCULATIONAHA.104.490599)
15. Cohen, J.B., Lotito, M.J., Trivedi, U.K., Denker, M.G., Cohen, D.L., Townsend, R.R. (2019) Cardiovascular Events and Mortality in White Coat Hypertension: A Systematic Review and Meta-analysis. *Annals of Internal Medicine*, 170(12): 853-862. DOI: [10.7326/M19-0223](https://doi.org/10.7326/M19-0223)
16. Kawamoto, K., Houlihan, C.A., Balas, E.A., Lobach, D.F. (2005) Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. *BMJ*, 330(7494): 765. DOI: [10.1136/bmj.38398.500764.8F](https://doi.org/10.1136/bmj.38398.500764.8F)
17. Bates, D.W., Kuperman, G.J., Wang, S., et al. (2003) Ten commandments for effective clinical decision support: making the practice of evidence-based medicine a reality. *Journal of the American Medical Informatics Association*, 10(6): 523-530. DOI: [10.1197/jamia.M1370](https://doi.org/10.1197/jamia.M1370)
18. Bright, T.J., Wong, A., Dhurjati, R., et al. (2012) Effect of clinical decision-support systems: a systematic review. *Annals of Internal Medicine*, 157(1): 29-43. DOI: [10.7326/0003-4819-157-1-201207030-00450](https://doi.org/10.7326/0003-4819-157-1-201207030-00450)
19. Roumie, C.L., Elasy, T.A., Greevy, R., et al. (2006) Improving blood pressure control through provider education, provider alerts, and patient education: a cluster randomized trial. *Annals of Internal Medicine*, 145(3): 165-175. DOI: [10.7326/0003-4819-145-3-200608010-00004](https://doi.org/10.7326/0003-4819-145-3-200608010-00004)
20. Green, B.B., Cook, A.J., Ralston, J.D., et al. (2008) Effectiveness of home blood pressure monitoring, Web communication, and pharmacist care on hypertension control: a randomized controlled trial. *JAMA*, 299(24): 2857-2867. DOI: [10.1001/jama.299.24.2857](https://doi.org/10.1001/jama.299.24.2857)
21. Margolis, K.L., Asche, S.E., Bergdall, A.R., et al. (2013) Effect of home blood pressure telemonitoring and pharmacist management on blood pressure control: a cluster randomized clinical trial. *JAMA*, 310(1): 46-56. DOI: [10.1001/jama.2013.6549](https://doi.org/10.1001/jama.2013.6549)
22. McManus, R.J., Mant, J., Haque, M.S., et al. (2014) Effect of self-monitoring and medication self-titration on systolic blood pressure in hypertensive patients at high risk of cardiovascular disease: the TASMIN-SR randomized clinical trial. *JAMA*, 312(8): 799-808. DOI: [10.1001/jama.2014.10057](https://doi.org/10.1001/jama.2014.10057)
23. Tang, P.C., Ralston, M., Arrigotti, M.F., Qureshi, L., Graham, J. (2007) Comparison of methodologies for calculating quality measures based on administrative data versus clinical data from an electronic health record system: implications for performance measures. *Journal of the American Medical Informatics Association*, 14(1): 10-15. DOI: [10.1197/jamia.M2198](https://doi.org/10.1197/jamia.M2198)

24. Persell, S.D., Dolan, N.C., Friesema, E.M., Thompson, J.A., Kaiser, D., Baker, D.W. (2010) Frequency of inappropriate medical exceptions to quality measures. *Annals of Internal Medicine*, 152(4): 225-231. DOI: [10.7326/0003-4819-152-4-201002160-00007](https://doi.org/10.7326/0003-4819-152-4-201002160-00007)
25. Persell, S.D., Kaiser, D., Dolan, N.C., et al. (2011) Changes in performance after implementation of a multifaceted electronic-health-record-based quality improvement system. *Medical Care*, 49(2): 117-125. DOI: [10.1097/MLR.0b013e318202913d](https://doi.org/10.1097/MLR.0b013e318202913d)
26. Muntner, P., Shimbo, D., Carey, R.M., et al. (2019) Measurement of Blood Pressure in Humans: A Scientific Statement From the American Heart Association. *Hypertension*, 73(5): e35-e66. DOI: [10.1161/HYP.0000000000000087](https://doi.org/10.1161/HYP.0000000000000087)
27. Powers, B.J., Olsen, M.K., Smith, V.A., Woolson, R.F., Bosworth, H.B., Oddone, E.Z. (2011) Measuring blood pressure for decision making and quality reporting: where and how many measures? *Annals of Internal Medicine*, 154(12): 781-788. DOI: [10.7326/0003-4819-154-12-201106210-00005](https://doi.org/10.7326/0003-4819-154-12-201106210-00005)
28. Rakotz, M.K., Townsend, R.R., Yang, J., et al. (2017) Medical Students and Measuring Blood Pressure: Results From the American Medical Association Blood Pressure Check Challenge. *Journal of Clinical Hypertension*, 19(6): 614-619. DOI: [10.1111/jch.13018](https://doi.org/10.1111/jch.13018)
29. Hripcsak, G., Albers, D.J. (2013) Next-generation phenotyping of electronic health records. *Journal of the American Medical Informatics Association*, 20(1): 117-121. DOI: [10.1136/amiajnl-2012-001145](https://doi.org/10.1136/amiajnl-2012-001145)
30. Goldstein, B.A., Navar, A.M., Pencina, M.J., Ioannidis, J.P.A. (2017) Opportunities and challenges in developing risk prediction models with electronic health record data: a systematic review. *Journal of the American Medical Informatics Association*, 24(1): 198-208. DOI: [10.1093/jamia/ocw042](https://doi.org/10.1093/jamia/ocw042)
31. Denny, J.C. (2012) Chapter 13: Mining electronic health records in the genomics era. *PLoS Computational Biology*, 8(12): e1002823. DOI: [10.1371/journal.pcbi.1002823](https://doi.org/10.1371/journal.pcbi.1002823)