The University Of Sheffield.

This is a repository copy of *Does workforce diversity, equity, and inclusion prevent patient safety incidents: a double machine learning approach*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/223930/

Version: Accepted Version

Article:

Wang, Y. orcid.org/0000-0003-1575-0245, Ji, J., Zhang, M. et al. (1 more author) (2025) Does workforce diversity, equity, and inclusion prevent patient safety incidents: a double machine learning approach. Journal of the Association for Information Systems, 26 (3). 6. ISSN 1558-3457

https://doi.org/10.17705/1jais.00944

© 2025 The Authors. Except as otherwise noted, this author-accepted version of a journal article published in Journal of the Association for Information Systems is made available via the University of Sheffield Research Publications and Copyright Policy under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution and reproduction in any medium, provided the original work is properly cited. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here: https://creativecommons.org/licenses/

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Does Workforce Diversity, Equity, and Inclusion Prevent Patient Safety Incidents: A Double Machine Learning Approach

Yichuan Wang
Sheffield University Management School, University of Sheffield, UK, yichuan.wang@sheffield.ac.uk

Jiao Ji
Sheffield University Management School, University of Sheffield, UK, jiao.ji@sheffield.ac.uk

Minhao Zhang
University of Bristol Business School, University of Bristol, UK, minhao.zhang@bristol.ac.uk

Xiaojun Wang
Birmingham Business School, University of Birmingham, UK, x.wang.18@bham.ac.uk

About the Author

Yichuan Wang is a Senior Lecturer/Associate Professor in Digital Marketing at Sheffield University Management School, the University of Sheffield. He holds a PhD in Business and Information Systems from the Raymond J. Harbert College of Business, Auburn University (USA). His research focuses on how digital technologies and information systems shape and influence a wide range of business practices. His research has appeared in journals such as *Journal of the Association for Information Systems*, *Social Science & Medicine*, *International Journal of Operations and Production Management*, *British Journal of Management*, and *Information & Management*. He currently serves as a Senior Editor for *The DATA BASE for Advances in Information Systems*. Since 2018, he has co-founded and organized the International Conference on Digital Health and Medical Analytics (DHA). He has led and collaborated with hospitals on various international projects that leverage healthcare information systems and business analytics to improve patient care quality.

Jiao Ji is an Assistant Professor of Finance at Sheffield University Management School, specializing in empirical corporate finance and governance with a focus on corporate risk, misconduct, and disclosure. Skilled in big data analytics, she uses textual analysis to explore corporate communications and social media content. Her research interests also span banking, innovations, environmental and social governance, and behavioral finance. Jiao's work has been featured in world-leading journals, and she contributes actively to scientific committees including the European Corporate Governance Institute, the British Accounting and Finance Association, and the Financial Management Association Europe. She also serves as an associate editor for Qualitative Research in Financial Markets and is a member of the audit research lab at the Centre for Research on Accounting and Finance in Context at the University of Sheffield, supported by a $500k grant from Luminate.

Minhao Zhang is an Associate Professor (equivalent to Reader) in Operations and Supply Chain Management at University of Bristol Business School. The themes throughout Minhao's current work are managing organizational risks through digitalization and data analytics. He is co-leading REWIRE (£11 million five-year Innovation and Knowledge Center on Semiconductor Technologies and Supply Chains) and REPHRAIN (£8.6 million three-year UKRI Research Centre of Excellence on Privacy, Harm Reduction, and Adversarial Influence Online). Thanks to various cross-disciplinary research collaborations, Minhao has published papers in the fields of operations management, information systems, and service research, earning top-cited and most-downloaded article awards from both the British Academy of Management and the European Academy of Management.

Xiaojun Wang is Professor of Operations and Supply Chain Management at Birmingham Business School, University of Birmingham. His current research includes low carbon manufacturing, supply chain management, healthcare operations, and machine learning & AI applications. He has published over 100 peer-reviewed papers in international renowned journals including Journal of Operations Management, Production and Operations Management, INFORMS Journal on Computing, European Journal of Operational Research, Omega, Decision Support Systems, Journal of Business Ethics, British Journal of Management, and International Journal of Production Economics. He has received funding for research from a range of funding bodies including NERC, ESRC, the Royal Society, the Newton Fund, and the National Natural Science Foundation of China (NSFC).

Does Workforce Diversity, Equity, and Inclusion Prevent Patient Safety Incidents: A Double Machine Learning Approach

Abstract

Workforce diversity, equity, and inclusion (DEI) are increasingly recognized as essential components in healthcare organizations. However, robust empirical research is scarce on how workforce DEI influences patient safety outcomes, particularly concerning the boundary conditions that may moderate this relationship. This study analyzes a longitudinal dataset from 2017 to 2021 that includes DEI metrics, staff-reported patient safety incidents, and employee feedback on DEI from Glassdoor and Indeed for 120 NHS Trusts in England's acute care sector. We examine workforce DEI through both its demographic and experiential dimensions to provide a comprehensive view. Employing a double machine learning approach, our findings indicate that a one-unit increase in workforce DEI scores is associated with a reduction of 8.108 patient safety incidents per 1000 admissions. Moreover, regions with greater patient racial diversity and healthcare organizations with lower complexity experience significantly greater benefits from DEI initiatives. This study provides healthcare policymakers and institutions with actionable insights to strategically tailor DEI initiatives and effectively improve patient safety.

Keywords: Diversity, equity, and inclusion (DEI), Patient safety, double machine learning, healthcare analytics

1 Introduction

Patient safety is a paramount indicator of healthcare quality, reflecting the standard of care within healthcare institutions (Nancy et al., 2016). An increased number of patient safety incidents may indicate shortcomings in care processes, thereby exposing patients to a heightened risk of harm (Tucker et al., 2020). The United Kingdom's National Health Service (NHS) has championed several initiatives, such as the implementation of Patient Safety Incident Response Framework to ameliorate patient safety. However, despite these measures, there has been a marked surge in reported incidents. Over 2.3 million patient safety incidents were reported by NHS organizations between April 2021 and March 2022, marking an 11.2% increase compared to the same period in 2020 (NHS England, 2022b). Among these reported incidents, here is increasing evidence that racial discrimination and social health disparities present formidable challenges to maintaining patient safety (Kapadia et al., 2022). Moreover, a staggering 76% of surveyed NHS professionals reported experiencing racism at work at least once in the past two years (British Medical Association [BMA], 2022). The chairman of the BMA emphasized the gravity of the situation, warning that persistent racism is causing many medical professionals to struggle with mental distress and feelings of isolation. This not only diminishes their overall well-being but also affects their ability to provide optimal care, potentially endangering patients (Nagesh, 2022).

Furthermore, systemic disparities affect the NHS's 250,000 Black and Minority Ethnic (BME) staff. They consistently face challenges ranging from recruitment and career progression biases to increased exposure to disciplinary actions and bullying. Discriminatory practices are evident in appointment rates: a White candidate who has been shortlisted is 1.61 times more likely to be selected than their BME counterpart (NHS England, 2022a). Such practices reveal a concerning trend where selection panels may be influenced by implicit biases, often favoring "people like us" or those who seem likely to "fit in." This not only deprives the

NHS of diverse talents but also limits its capacity to address the unique health challenges and needs of a multicultural patient demographic (Molloy, 2022).

In addressing patient safety concerns, the role of workforce diversity, equity, and inclusion (DEI) is becoming increasingly apparent. Diversity refers to the representation of physical and socio-cultural differences among individuals in organizations (Arsel et al., 2022; Rotenstein et al., 2021). Equity refers to treating individuals in organizations fairly by offering equal opportunities and desirable outcomes (Arsel et al., 2022). Inclusion refers to creating an environment in which individuals feel respected, accepted, valued, belonging, and incorporated, and able to participate in decision-making processes (Arsel et al., 2022; Romansky et al., 2021). Workforce DEI can be viewed as a comprehensive measure of an organization's commitment to these principles, with a particular focus on traditionally marginalized groups. For example, the UK government has prioritized reducing inequalities by integrating DEI into the Patient Safety Incident Response Framework (NHS England, 2022b), and the National Institutes of Health (NIH) has launched the UNITE initiative to tackle structural racism and health disparities (Boulware et al., 2022a). However, healthcare advocates sometimes refer to the designation of "a titular diversity-equity-inclusion chief" (Grubbs, 2020, p. e25) or merely to "tick box" exercises (NHS East of England, 2021) as a means of promoting DEI in the workforce. Consequently, the effectiveness of workforce DEI in enhancing patient safety remains ambiguous and requires deeper examination.

Historically, specific DEI facets, from gender diversity (e.g., Chang et al., 2020; Cumming et al., 2015) to religious and sexual orientation in recruitment processes (Acquisti & Fong, 2020) and further to racial and income representation in branding practices (Park et al., 2023) have been explored in the management literature. The medical literature documents that DEI efforts are increasingly being integrated into healthcare delivery, such as diversifying the research workforce in clinical trials (Boulware et al., 2022b) and promoting shared responsibility for diversity (Rotenstein et al., 2021). However, there is a lack of empirical evidence showing that efforts to promote workforce DEI lead to a significant reduction in patient safety incidents. We address this gap by examining 1) *Can workforce DEI serve as a leading indicator of patient safety incident reduction?* 2) *If so, what specific characteristics of workforce DEI are more pivotal in predicting patient safety incident reduction? 3*) *What boundary conditions might moderate the relationship between workforce DEI and patient safety incidents?*

This study embarks on a pioneering exploration of workforce DEI, making a unique contribution to the field of information systems (IS). Although IS literature has emphasized the importance of health data inclusiveness in mitigating algorithmic biases (e.g., Agarwal et al., 2020; Bardhan et al., 2020) and highlighted the role of health information technology (HIT) in reducing healthcare access inequalities (Tong et al., 2022), the development of DEI within the workforce remains an emerging societal challenge. This challenge is closely linked to employee experiences (Rotenstein et al., 2021) and patient outcomes (Simsekler & Qazi, 2022), yet it has been largely underexplored by IS scholars.

In response to Marabelli and Chan's (2024) call for more IS research addressing DEI, this paper aims to fill that gap within the context of healthcare analytics, as advocated by Bird et al. (2023). We combine double machine learning (DML), an innovative integration of econometrics and machine learning, with the analysis of unstructured data to reveal the impact of workforce DEI on patient safety outcomes. By harnessing natural language processing (NLP) to extract actionable insights from disparate data sources, including social media, our approach not only highlights the importance of advanced analytics in identifying trends within the healthcare sector but also captures healthcare professionals' authentic voices. This work exemplifies the potential of novel methods and diverse data utilization to uncover critical patterns and insights, ultimately contributing to improved patient safety outcomes.

Our investigation not only confirms the profound impact of workforce DEI on patient safety outcomes but also identifies which specific DEI characteristics are most influential. As healthcare organizations increasingly embrace predictive modeling, the potential of methodologies such as DML becomes evident. Our study demonstrates that the DML approach is exceptionally adept at estimating the causal effects of DEI initiatives on patient safety. By integrating such methodologies, healthcare institutions can more accurately anticipate the implications of their DEI efforts and tailor these initiatives for maximum impact. Furthermore, by identifying key contextual factors (e.g., regional racial diversity and organizational complexity) hat moderate the relationship between workforce DEI and patient safety incidents, we offer a more holistic understanding that challenges traditional, one-size-fits-all paradigms.

2 Theoretical Background

2.1 Workforce Diversity, Equity, and Inclusion

The current landscape of DEI measures varies in scope and methodology across studies. Our review of prior DEI measures (See Table 1) reveals several noticeable gaps. ***First***, many studies rely on single-dimensional DEI observations, a limitation that stems both from the scarcity of available DEI information and from challenges in harnessing diverse data sources. For instance, Acquisti and Fong (2020) focused solely on religious affiliation and sexual orientation, while Robert et al. (2018) and Park et al. (2023) examined racial dynamics, with the latter further including income representativeness. Although Simsekler and Qazi (2022) adopted a broader approach by incorporating fairness in career progression and discrimination experiences, their measures still remain limited in scope. A more holistic, multidimensional view of DEI is seen in studies by Dube and Zhu (2021) and Li et al. (2024). The former integrated a range workplace dimensions, from employee diversity to health and safety. The latter concentrated on DEI announcements within organizations. The inherent risk of single-dimensional measures is that they tend to overgeneralize, potentially overshadowing specific subgroups within broader categories. This evolution toward multidimensionality reflects the complexity of DEI and highlights the inadequacy of unidimensional measures. Therefore, there is a clear need for efforts that uniquely measure workforce DEI while incorporating perspectives from multiple stakeholders.

Second, panel data analysis emerges as a powerful tool for addressing the shortcomings of previous research methodologies (e.g., Li et al., 2024; Park et al., 2023). A primary challenge in DEI research has been omitted variable bias, as highlighted in Simsekler and Qazi's (2022) study. This bias occurs when pertinent variables are left out of the model, leading to inaccurate or misleading results. Panel data analysis remedies this issue by capturing information over time and across different subjects, allowing researchers to identify and control for unobserved heterogeneity, which includes latent variables or unseen factors that may influence the relationship between studied variables (Zyphur et al., 2020 Traditional cross-sectional DEI studies often miss these nuances, potentially resulting in spurious relationships and mistaken inferences. In addition, the richness of panel data facilitates multilevel modeling. This is particularly relevant to DEI research, as DEI dynamics operate at multiple levels such as individual, organizational, regional, and national. Individual experiences are shaped not only by organizational culture and dynamics but also by overarching national policies.

Third, much of the prior DEI research has predominantly used traditional methodologies, often relying on primary data collected through internal employee surveys (Simsekler & Qazi, 2022) and experiments (Acquisti and Fong, 2020; Robert et al., 2018). Although internal DEI surveys are valuable, they sometimes face limitations such as concerns about anonymity and potential repercussions for negative feedback (Simsekler & Qazi, 2022). Our study fills a unique niche by exploring healthcare professionals' online reviews on social media platforms such as Glassdoor and Indeed using text mining analytics. Online platforms provide a forum where employees can express their authentic voices and sentiments. In contrast to traditional surveys, where respondents may be influenced by social desirability bias, company loyalty, or fear of retribution, online reviews offer a setting in which employees tend to speak more candidly. In addition, online reviews offer a temporal perspective on DEI perceptions, capturing shifts and trends over time that one-off surveys or experiments might miss. Empirical evidence demonstrates the predictive power of these reviews, as they have been linked to operational performance (Huang et al., 2020), corporate disclosures (Hales et al., 2018), CEO dismissals (Wang et al., 2022a), and even instances of corporate misconduct (Campbell & Shang, 2022). By harnessing the capabilities of text analytics, our methodology moves beyond anecdotal evidence to provide quantifiable insights from these online reviews. In doing so, we not only highlight real-time DEI perceptions but also underscore the potential of digital footprints to offer deeper and more nuanced insights into the complexities of DEI in contemporary workplaces.

Table 1. A summary of prior DEI measures

Studies	Focus of DEI	DEI measures	Multidimensional measure developed	Text mining used for developing measures	Multiple data sources	Longitudinal data used
Acquisti and Fong (2020)	Religious affiliation and sexual orientation	A between-subjects design with four treatment conditions: religious affiliation (a Christian versus a Muslim male) and sexual orientation (a gay versus a straight male)				
Dube and Zhu (2021)	Employee diversity, compensation and benefits, work-life balance, union relations, and health and safety	Assess firms' workplace practices using the MSCI ESG KLD STATS database between 2003 and 2018, considering both strengths and concerns related to employee relations and diversity	√		√	√
Li et al. (2023)	DEI commitment	Use Factiva to search DEI-related announcements	√			√

		from 2014 to 2022 using specific DEI keywords				
Park et al. (2023)	Customer DEI as an outcome of interest	Use household-level transaction data to estimate customer DEI outcomes, which include racial and income representativeness			√	√
Robert et al. (2018)	Racial and gender diversity	Conduct a laboratory experiment involving 46 teams, predominantly composed of individuals who self-identified as Caucasian and Asian, both men and women	√			
Simsekler and Qazi (2022)	DEI perception	DEI survey containing four items such as fairness of career progression, discrimination at work, and adjustment				
Current study	Race diversity, gender equity, and DEI experience	A longitudinal dataset spanning 2017-2021 provides DEI metrics, patient safety incidents, and Glassdoor/Indeed DEI feedback for 120 acute care NHS Trusts in England.	√	√	√	√

2.2 Workforce DEI in Predicting Patient Safety Incident Reduction

Patient safety competencies among healthcare professionals are influenced by individual characteristics, such as professional knowledge (Okuyama et al., 2011), technical characteristics, including information accessibility (Hydari et al., 2019; Ranganathan et al., 2004), and organizational characteristics, such as training and education (Ginsburg et al., 2012). At the core of these competencies is the patient safety culture within healthcare settings, which is deeply affected by professionals' sense of belonging to their organization (Mael and Ashforth, 1992, p.104). Hu and Casey (2021) emphasize that this sense of belonging is crucial for ensuring patient safety. Healthcare professionals, particularly those from traditionally underrepresented, underserved, and marginalized communities, may face unequal treatment and discrimination in the workplace. They can encounter bias and prejudice from both coworkers and patients, which may lead to feelings of detachment from the organization. This detachment can result in a reduced commitment to organizational goals, such as reporting safety concerns and potential risks, and ultimately foster a hostile work environment that undermines their ability to deliver effective care.

Addressing these concerns underscores the need for workforce DEI initiatives that go beyond more demographic representation (Guillaume et al., 2017) to rectify systemic disparities and biases, thereby fostering a more inclusive, respectful, and equitable work environment (Simsekler & Qazi, 2022). ***Demographic DEI*** focuses on visible or quantifiable characteristics within workgroups, such as gender, race, age, and ethnicity, and emphasizes the equitable treatment of individuals from diverse backgrounds to mirror the broader society. Van Dijk et al. (2012) note that demographic diversity is linked to enhanced workgroup performance, particularly in tasks that require innovation or involve task-relevant knowledge. In the marketing literature, effective diversity management has been shown to not only improve customer satisfaction but also enhance sales performance when store-unit racial diversity aligns with community diversity (Patel & Feng, 2021; Park et al., 2023).

Experiential DEI refers to the lived experiences, backgrounds, and perspectives that individuals bring to an organization (Simsekler & Qazi, 2022; Randel et al., 2016), complementing demographic DEI. Sabharwal (2014) contends that for enhanced performance, organizations should focus not only on diversity but also on inclusion, valuing employee opinions and boosting self-esteem. In healthcare settings, the complexity of tasks often requires a range of experiences, which leads to a more creative problem-solving environment and improves the ability to address complex patient safety challenges. For instance, when healthcare organizations prioritize racial diversity, the resulting workforce of varied groups can help curb in-group favoritism and promote more unbiased, holistic decision-making that is crucial for reducing patient safety incidents. Grounded in social categorization and information decision-making perspectives, workplace diversity has been linked to performance-related variables such as organizational performance, workgroup performance, innovation, and individual in-role and extra-role performances (Joshi et al., 2011; Guillaume et al., 2017). Self-categorization theory explains how individuals perceive and categorize themselves and others based on similarities and differences between their own group and other groups (Hogg & Terry, 2020). In IS literature, several studies have applied self-categorization theory to examine group behaviors in system development, although the results have been inconsistent. Lee and Xia (2010) suggest that within software development, diversity can be a double-edged sword. While it offers varied perspectives, it might also introduce communication hurdles that could weaken team cohesion and undermine agility. Conversely, Wang et al. (2022b) argue that fostering alignment and close relationships among groups can strengthen organizational identification among employees, which is instrumental in retaining talent and effectively reducing turnover rates.

Moreover, the gender pay gap in medicine has been a longstanding concern (Gottlieb et al., 2021). Addressing and rectifying this disparity can help reduce pronounced in-group and out-group distinctions based on gender. Ensuring equitable compensation across all genders not only breaks down divisive barriers created by monetary discrimination but also enhances team cohesion. Recognizing every healthcare professional's worth through equal pay is a testament to respect, fairness, and acknowledgment (Hoff, 2021), fostering a profound sense of belonging and allegiance to the institution. Thus, we contend that this heightened commitment motivates healthcare professionals to exhibit enhanced diligence in their duties, especially in areas crucial to reducing patient safety incidents. A tangible example of the benefits of addressing the gender pay gap can be seen in Denmark's 2006 legislative change, which mandated pay transparency by requiring firms to provide gender-disaggregated wage statistics (Bennedsen et al., 2022). This legislation reduced the gender pay gap by 2 percentage points, a 13% decline relative to the pre-legislation mean, primarily by moderating wage growth for male employees.

In summary, the relationship between workforce DEI and the reduction of patient safety incidents is multifaceted. By fostering both demographic and experiential DEI, healthcare organizations can create a synergistic environment that reduces patient safety incidents. Thus, we hypothesize:

Hypothesis 1a-c (H1a-c)*: Healthcare organizations that a) exhibit greater race diversity, b) have narrower gender pay gaps, and c) provide a positive DEI experience will witness a reduction in patient safety incidents.*

2.3 What Moderates the Effects of Workplace DEI on Patient Safety Incidents

2.3.1 DEI information transparency

Healthcare organizations that effectively implement DEI initiatives may reduce the likelihood of patient safety incidents related to workforce issues. However, merely having these initiatives is not sufficient. Their effectiveness depends on transparent implementation, ongoing monitoring, and continuous improvement. As defined by Granados and Gupta (2013), information transparency acts as an information strategy to "*selectively disclose information outside the boundaries of the firm, to buyers, suppliers, competitors, and other third parties like governments and local communities*" (p.638).

Previous research has explored the effects of company information transparency. For instance, Mas (2016) found that enforced transparency of executive earnings led to higher average CEO compensation relative to other top-earning executives. Bennedsen et al. (2022) demonstrated that pay transparency reduced the gender pay gap by 13 percent without impacting firm profitability, mainly by slowing wage growth for male employees. In the academic realm, Obloj and Zenger (2022) provided empirical evidence over two decades that pay transparency significantly narrowed the gender pay gap in academia. More recently, new laws in New York City and California mandated the inclusion of pay ranges in job listings in 2023 to reduce ambiguity during compensation negotiations (Tang, 2024).

In contrast, Mas (2017) observed that after a 2010 mandate in California required the online posting of municipal salaries, top managers experienced an average compensation decrease of 7 percent and a 75 percent increase in quit rates. Chen et al. (2022) also noted that full transparency regarding granular compensation details may not be necessary; simply disclosing average compensation among peers can be enough to enhance worker motivation and performance. These mixed findings raise intriguing questions about the broader implications of information transparency beyond the corporate sector. In healthcare, the intrinsic values and operational imperatives, centered on patient welfare, ethical practice, and equitable treatment, may create a more receptive environment for transparency. Transparency in DEI initiatives could strengthen patient outcomes by exposing underlying biases and prompting proactive change. DEI in healthcare is not just an administrative or human resource concern; it fundamentally affects the institution's ethos and operational efficiency (Stanford, 2020).

Emphasizing transparency in DEI initiatives is crucial because it moderates the relationship between workforce DEI and patient safety outcomes. When information about DEI efforts is transparent, patients perceive a genuine institutional commitment to DEI. This heightened perception of commitment can increase patient trust, which in turn leads to better adherence to medical recommendations, stronger patient-provider relationships, and improved health outcomes. Conversely, if DEI efforts are perceived as opaque or superficial due to limited information transparency, this trust may erode, potentially weakening the positive impact of workforce DEI on reducing patient safety incidents. Therefore, we propose:

Hypothesis 2 (H2)*: For healthcare organizations with greater DEI information transparency, the negative relationship between workforce DEI and patient safety incidents is stronger than in organizations with limited DEI information transparency.*

2.3.2 Organizational complexity

Complexity theory describes complexity as a "structural variable that characterizes both organizations and their environments" (Anderson, 1999, p. 216). Complexity arises because multiple subsystems operate simultaneously within an organization. Research has indicated that complexity increases the likelihood of project failure, such as in IT projects (Xin and Choudhary, 2019). The reasons for implementation failure can be linked to organizational, technical, social, and industrial environmental factors, with organizational complexity being one of the major causes. Depending on the organizational size and structure, factors that contribute to organizational complexity may affect the effectiveness of DEI initiatives. When implementing DEI, it is crucial to consider organizational complexity as a contextual variable (Tang, 2024) to effectively formulate DEI strategies that yield positive patient outcomes under various organizational conditions.

However, effectively advocating for DEI initiatives requires a significant investment of these operational resources (Jackson, 2023). In highly complex organizational settings, insufficient resource allocation to DEI initiatives may result from competing priorities. For example, in a large hospital system, DEI initiatives aimed at enhancing patient safety through equitable care could be underfunded due to urgent reallocations toward emergency services and medical equipment upgrades. Consequently, the intended improvements in cultural competence and equitable patient treatment may not be fully achieved, thereby impacting the quality of patient care.

Furthermore, complex organizations often have entrenched cultures and subcultures that resist new initiatives, especially those that challenge long-standing norms and practices, such as DEI initiatives (Hellerstedt et al., 2024). This resistance can be particularly pronounced in environments where diverse and sometimes conflicting subcultural identities and values exist. In a study of multinational enterprises, Ferner et al. (2005) observed that subsidiary managers used their power to resist the adoption of new diversity policies from headquarters, highlighting a significant challenge in implementing DEI initiatives in complex organizational settings. This resistance, stemming from organizational complexity, can impede the adoption and effective implementation of DEI principles, thereby diminishing their impact on organizational practices and outcomes. We thus hypothesize that:

> ***Hypothesis 3 (H3)****: In healthcare organizations characterized by higher levels of organizational complexity, the effectiveness of workforce DEI initiatives in reducing patient safety incidents is diminished compared to organizations with lower complexity.*

2.3.3 Regional patient race diversity

The Categorization-Elaboration Model (CEM) proposed by van Knippenberg et al. (2004) defines diversity as variations among individuals on any attribute that creates perceptions of dissimilarity. According to the CEM, the outcomes of diversity, whether advantageous or disadvantageous, are moderated by factors that highlight demographic differences and those that affect information elaboration. The principle of comparative fit suggests that social categorizations become more salient when pronounced intergroup differences overshadow intra-group similarities (van Knippenberg et al., 2004). When applied to regions with significant patient racial diversity, it becomes imperative that the healthcare workforce reflects a similar level of diversity. This alignment between workforce DEI and patient demographics is crucial for several reasons.

First, as noted by Earley and Mosakowski (2000), in groups comprising multiple nationalities or racial backgrounds, the tendency to categorize members based on a single dimension, such as nationality, is reduced. Similarly, in a healthcare setting with extensive racial diversity among patients, a diverse workforce helps prevent oversimplified classifications based solely on race. **Second**, a workforce that mirrors the racial diversity of its patient demographic can prevent an "us-versus-them" mentality (Guillaume et al., 2017; van Knippenberg et al., 2004), which risks compromising the quality of patient care. ***Finally***, a racially diverse patient population brings a range of healthcare needs, cultural considerations, and communication nuances. A workforce whose diversity aligns with that of its patients is better positioned to understand these complexities, fostering trust and proactively addressing patient safety issues. In light of these insights, we posit the following hypothesis:

> ***Hypothesis 4 (H4):*** *In regions with higher patient racial diversity, the negative relationship between workforce DEI and patient safety incidents is more pronounced compared to regions with lower patient racial diversity.*

3 Research Method

3.1 Data and Measures

To advance our theorizing, we have compiled a unique dataset that tracks extensive longitudinal panel data on NHS workforce DEI, NHS staff-reported patient safety incidents, and employee DEI feedback from Glassdoor and Indeed across 120 NHS Trusts in England's acute care sector over a five-year period (2017–2021). Our final dataset comprises over 90% of acute care NHS trusts in England, totaling 600 observations. The data were collected from multiple sources at the regional, organizational, and individual levels. Primarily, we used publicly available official NHS data provided by NHS Digital. To complement this dataset, we collected information on the gender pay gap and DEI experiences from the GOV.UK website, and employee reviews from social media platforms (Glassdoor and Indeed). Figure 1 illustrates the research roadmap. In particular, our focus is on investigating the role of workforce DEI in predicting patient safety incidents and exploring the boundary conditions that moderate the DEI-patient safety incident link.

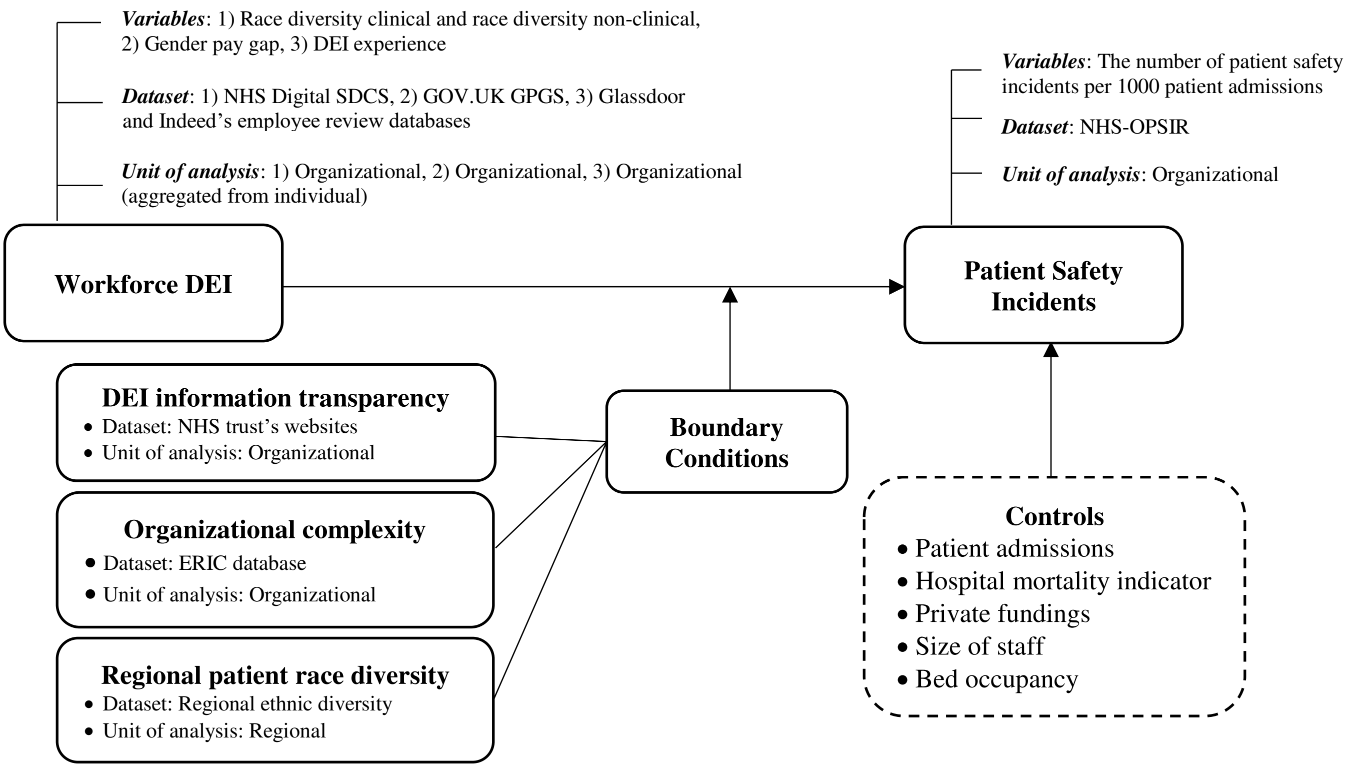


Figure 1. Research Road Map

3.1.1 Patient safety incidents

Within the NHS, patient safety is primarily assessed through two avenues: patient-reported measures and staff-reported incidents. Although patient-reported feedback provides crucial insights into individual experiences and perceived healthcare quality, these perspectives can be influenced by regional cultural nuances, public perceptions of a hospital, staff responses to complaints, and individual beliefs about the national health service (Gillespie & Reader, 2023). In contrast, staff-reported incidents offer a more comprehensive view. As they are directly involved in care delivery, staff members are better positioned to identify, understand, and document safety breaches ranging from minor infractions to severe events that cause long-term harm or death. Given this context, our study relies on staff-reported patient safety incidents as the dependent variable. This measure is obtained from the UK National Reporting and Learning System and represents the number of staff-reported patient safety incidents per 1,000 patient admissions.

3.1.2 Demographic aspect of workforce DEI

Selecting a single metric to encapsulate the NHS trust's DEI performance across race and gender dimensions is notably complex, given the multitude of potential indicators[1]. To address this challenge, we employ the LASSO (Least Absolute Shrinkage and Selection Operator) approach to identify the most impactful race and gender metrics within our constrained sample size. LASSO is a key technique in modern statistics and machine learning, offering a robust approach for handling high-dimensional data through variable selection and regularization, thereby enhancing model simplicity and interpretability (e.g., Mullainathan & Spiess, 2017). LASSO minimizes the sum of squared residuals while applying a penalty (λ) to the absolute magnitude of the coefficient estimates. As λ increases, more coefficients are reduced to zero and effectively eliminated, reducing variance at the expense of increased bias, a trade-off that ultimately enhances the predictive accuracy of the model. The optimal penalty level is determined using the Extended Bayesian Information Criterion (EBIC) (Chen & Chen, 2008). Recognizing that race and gender capture distinct aspects of DEI in the England NHS sector,

[1] We gather workforce race diversity data from the NHS Digital Strategic Data Collection Service (SDCS) to assess demographic diversity in DEI. The race diversity data encompass diversity levels for both the clinical and non-clinical workforce, spanning support, meddle, senior, and very senior managers positions within the NHS trust. On the other hand, we obtained gender pay gap data from the GOV.UK website.

our objective is to leverage LASSO to isolate the most influential predictors from each category. This process enables us to construct a more precise and meaningful DEI index, which serves as our main independent variable. These methods ensure a comprehensive and robust assessment of an NHS trust's DEI level, accounting for both race and gender dimensions.

For race diversity information, we examined all available data from NHS Digital. We collected three groups of indicators that may represent race-related DEI: 1) race-demographic diversity for clinical staff (*support rank staff BME ratio, middle rank staff BME ratio, senior rank staff BME ratio, very senior manager rank staff BME ratio*), 2) race-demographic diversity for non-clinical staff (*support rank staff BME ratio, middle rank staff BME ratio, senior rank staff BME ratio, very senior manager rank staff BME ratio*), and 3) racial differences in work experiences reported in NHS Digital (*recruitment and selection, formal disciplinary, training and development, equal opportunities*).

Given the large number of potential predictors and our relatively small sample size, there is a high risk that an OLS model would overfit the data. Overfitting could result in identifying predictors that are significant only by chance (false positives) and that perform poorly on new data. Moreover, we do not know the true model or which specific race DEI regressors are most important. To address this, we introduced all these variables into our model (a full list is provided in Appendix A) and conducted a post-LASSO analysis, following the two-step procedure outlined by Belloni et al. (2014a, 2014b). In the first step, the LASSO method selects the variables that best predict the outcome, Patient Safety Incidents. In the second step, we apply standard OLS regression using only the variables chosen in the first step. The post-LASSO analysis selected the following variables for assessing race diversity: *Race Diversity Clinical_1 (%)*, which measures the percentage of non-white clinical staff at a senior level; *Race Diversity Clinical_2 (%)*, which quantifies the percentage of non-white clinical staff at a middle level; and *Race Diversity Non-Clinical (%)*, which represents the percentage of non-white non-clinical staff at a support level.

We use the UK national gender pay gap data (https://gender-pay-gap.service.gov.uk/). Since 2017, NHS trusts have been required to disclose information on the gender pay gap, an essential component of employee DEI. DEI. Mandatory gender pay gap (GPG) reporting was introduced in the UK in 2017 with the aim of narrowing and eventually eliminating the pay differential between men and women. The gender pay gap data includes information on the ratio of females across various staff levels, as well as pay disparities in salaries and bonuses. Notably, the NHS workforce is comprised of 77 percent women. Recently, there has been heightened emphasis on the gender pay gap in the NHS, driven by the mandatory annual publication of data for all large employers (over 250 employees). We collected various variables such as the gender pay gap in hourly pay, the gender pay gap in bonuses, and gender diversity across different quartiles (Lower Quartile, Lower Middle Quartile, Upper Middle Quartile, and Top Quartile). The proxy selected by the LASSO analysis is the Gender Pay Gap (%) variable, which represents the median difference in salaries between female and male employees.

After completing the LASSO process, we identified the primary indicators that depict the demographic facets of workforce DEI: Race Diversity Clinical_1, Race Diversity Clinical_2, Race Diversity Non-clinical, and Gender Pay Gap. We did not emphasize the theoretical significance of each variable selected by LASSO; rather, we used LASSO to pinpoint those variables in each category that have the strongest predictive power for the dependent variable. This approach yields better out-of-sample predictions than traditional variable selection methods, such as stepwise regression, which tend to be more prone to overfitting within the sample. We anticipated that this process would generate at least one robust proxy for each category.

3.1.3 Experiential aspect of workforce DEI

In addition to the demographic DEI, employee experiences are crucial for understanding DEI perspectives (Guillaume et al., 2017). To capture these experiences, we analyzed 15,133 employee reviews from Glassdoor and Indeed for 120 NHS Trusts spanning 2017 to 2021. This approach complements internal staff satisfaction surveys, which may not fully capture honest DEI opinions due to potential reluctance in sharing feedback openly. Notably, the NHS began collecting detailed DEI information in 2020, aligning with the government's Diversity and Inclusion Strategy for 2019–2023. Using deep learning-enabled analytics, we extracted DEI-related narratives from these social media reviews, capturing the authentic voices of employees. This method provides valuable insights into their experiences and perceptions of DEI within the NHS, allowing us to assess progress and changes over time. Glassdoor and Indeed employ structured formats that enable employees to evaluate their organizations across various dimensions. Reviews on these platforms are segmented into pros and cons, as illustrated in Figure B-1 in Appendix B, which facilitates sentiment analysis, especially regarding DEI. By examining these segments separately, researchers can better gauge both the successes and challenges of DEI initiatives within organizations.

The sentiment of terms used in these sections is essential, as their meaning can vary significantly by context. For example, when diversity training is mentioned in the pros section, it typically signals approval of the organization's DEI efforts. In contrast, in the cons section, the same term might indicate dissatisfaction with the effectiveness of these initiatives. Campbell and Shang (2022) argue that analyzing the cons section alone can significantly predict corporate misconduct. However, our study focuses exclusively on the pros section to assess positive employee experiences related to DEI, thereby highlighting and measuring the constructive aspects of these efforts.

To prepare the textual data for analysis, we performed extensive preprocessing. All text was converted to lowercase to ensure uniformity and eliminate inconsistencies due to capitalization. We then removed common English stopwords using the Natural Language Toolkit (NLTK) stopwords list, focusing on substantive words that carry meaningful content. Punctuation and non-alphanumeric characters were stripped from the text to eliminate noise, and the cleaned text was tokenized into individual words using NLTK's tokenizer, laying the groundwork for further analysis.

To capture meaningful phrases and contextual nuances, we generated unigrams (single words), bigrams (two-word phrases), and trigrams (three-word phrases) using Gensim's Phrases and Phraser classes. Including n-grams is essential because DEI-related concepts often involve multi-word expressions such as "equal opportunity" or "diverse workforce." This step allows us to identify these phrases that might not be evident when considering unigrams alone (Sautner et al., 2023). We set the minimum count for phrases at five occurrences in the corpus to ensure statistical significance and adjusted the threshold parameter to balance the inclusion of meaningful phrases without introducing spurious combinations. This preprocessing strategy aligns with best practices in natural language processing and text analytics, enabling the effective identification of relevant DEI terms in subsequent steps.

Dictionary Expansion Using Word Embeddings. With the pre-processed data in hand, we proceeded to expand our DEI dictionary to capture a broader range of relevant terms used by employees. Our starting point was the set of diversity and inclusion terms frequently mentioned in the UK Policy paper "Diversity and Inclusion Strategy 2019 to 2023." This document highlights key terms such as fairness, respect for difference, opportunity equality, treatment equality, inclusive environment, belongingness, safety, and inspired engagement, each serving as a core "DEI keyword" that expresses fundamental workforce values. We trained a Word2Vec model on the corpus using the Skip-gram architecture, which captures semantic relationships between words based on their contextual usage and generates word embeddings that reflect semantic similarity (Li et al., 2021). We set the hyperparameters with a vector size of 100 to capture nuanced relationships, a window size of five words to consider the surrounding context, and a minimum word frequency threshold of five. The model was trained over ten iterations to ensure convergence and stability.

Using the trained Word2Vec model, we computed cosine similarity scores between our seed words and other words in the corpus. Cosine similarity quantifies semantic similarity by assessing the cosine of the angle between two vectors in the embedding space. Following methodologies from prior studies (Li et al., 2021), we selected the top 500 words with the highest similarity scores relative to the average vector of the seed words. This approach allowed us to capture a broad spectrum of DEI-related language while maintaining relevance. A manual review was then conducted to verify the contextual appropriateness of the expanded terms, resulting in the removal of some unrelated words. Our final expanded dictionary included 475 DEI-related words.

Matching DEI Terms and Calculating DEI Index. At the hospital-year level, we aggregated the data to analyze temporal trends and organizational differences in DEI discussions. Given the significant impact of valence in interpreting DEI discussions, we specifically focused on positive aspects (i.e. Pros section). We organized the preprocessed reviews by hospital codes and years. For each hospital-year grouping, we calculated the frequency of each DEI term from the expanded dictionary in the Pros sections. Term frequencies provide a straightforward measure of how often DEI topics are mentioned. However, to account for the importance of terms within and across documents, we also calculated Term Frequency-Inverse Document Frequency (TF-IDF) scores. TF-IDF weighting reduces the impact of commonly used words and emphasizes terms that are significant in specific documents (Salton and Buckley, 1988). In our study, using TF-IDF weights highlights the significance of terms beyond mere frequency, capturing the importance of DEI discussions within the context of the entire corpus. We calculated our DEI index as the proportion of DEI words, weighted by their TF-IDF score in the Pros section, divided by the total number of words:

$$DEI\ index_{i,t} = \frac{\sum \omega \in DEIWords(Count_{i,t,\omega}^{Pors} \times TFIDF_{\omega})}{TotalWords_{i,t}}$$

Where ω denotes DEI-related words from our previously expanded dictionary, which consists of 475 words. $Count_{i,t,\omega}^{Pros}$ represents the frequency of word ω in the *Pros* section for hospital i in year t. $TFIDF_{\omega}$ is the TF-IDF weight of word ω, reflecting its importance across the entire corpus. To enhance interpretability and prevent the index from being excessively small, we multiplied the normalized values by a scaling factor of 1,000, effectively expressing the index per 1,000 words. This scaling practice aligns with established methodologies in textual analysis research, facilitating comparisons across different contexts and studies (Sautner et al., 2023).

Table B-1 (see Appendix B) presents the top 30 DEI terms extracted from employee reviews, separately listed for the *Pros* sections, analyzed by both raw frequency and TF-IDF weighting. This table offers a clear visualization of the most prominent DEI-related language used by employees in positive contexts. In the Pros section, frequently occurring words such as "diversity," "supported," and "valued" indicate positive acknowledgment of the organization's DEI efforts. The inclusion of TF-IDF weighted terms highlights words that, while they may not occur frequently, hold significant importance within specific reviews, emphasizing nuanced sentiments that mere frequency counts might overlook.

3.1.4 Principal components analysis (PCA) of workforce DEI

We construct our composite workforce DEI index using PCA, a method that maximizes the common variation captured by the index. PCA assigns weights to each dimension to optimize the shared variance among the workforce DEI measures. Appendix C reports the results of our PCA analysis. Panel A shows that three of our four workforce DEI measures load on the first principal component in a way that suggests lower workforce DEI (or greater homogeneity) for higher component values. Panel B indicates that this first principal component accounts for approximately 54.8% of the common variation across the four measures and has an eigenvalue substantially higher than one. Therefore, this component provides a reasonable summary measure of the common variation in the different dimensions of workforce DEI, which we use as our baseline index.

3.1.5 Boundary conditions

DEI Information Transparency. To measure DEI information transparency, we conducted a thorough examination of each NHS Trust's website to assess its public disclosure of workforce DEI information. We observed a range of DEI discourse practices: while some NHS trusts comprehensively disclosed data on ethnic diversity, gender equity, and other recommended workforce DEI metrics, others limited their disclosures to DEI data. Consequently, we define DEI information transparency as a binary indicator that equals one if an NHS trust fully disclosed all DEI information. Appendix D provides examples illustrating full transparency (coded as 1) and limited or no disclosure (coded as 0).

Organizational Complexity. We employ a binary variable to represent the complexity of each NHS trust. This variable is assigned a value of one if the trust is classified as large or teaching in the Estates Return Information Collection (ERIC) dataset, and zero if it is categorized as small, medium, or multi-service. This approach helps differentiate the levels of complexity among NHS trusts in our study.

Regional Patient Race Diversity is a binary variable set to one if the associated NHS trust is located in a region of England where the non-white population exceeds the national median non-white population level, based on 2021 Census data[2]. The England and Wales Census, conducted by the Office for National Statistics (ONS), has been carried out every ten years since 1801, capturing detailed information about individuals and households. The 2021 Census achieved a 97% response rate among the usually resident population of England and Wales, surpassing its target. By considering the entire population rather than a sample, the Census offers a comprehensive snapshot with minimal margin for error. Its accuracy is further enhanced through rigorous quality assurance processes, including cross-referencing with a wide array of alternative data sources.

3.1.6 Controls

Patient Admissions reflects patient admissions per 1,000 patients, sourced from the Hospital Episode Statistics database. *Hospital Mortality Indicator (SHMI)* Ranking assesses mortality rates in NHS hospital trusts in England. It compares actual patient deaths to expected deaths based on national averages, resulting in three scores: (1) High Mortality, (2) Average Mortality, and (3) Lower-than-Expected Mortality. *Funding* is measured as the natural log of private funding in a trust for each year, based on data from ERIC. Private investment in NHS trusts can significantly support and enhance the healthcare system, including infrastructure development and facility improvements, ultimately leading to better services and treatments for patients. *Size of Staff* is measured as the natural log of the total number of staff in a hospital trust, derived from the National Workforce Data Set. *Bed Occupancy* is the percentage of occupied beds in a hospital trust, obtained from NHS Bed Availability and Occupancy Data. This measure serves as a proxy for the trust's operational activity.

[2] High Ethnicity Diversity Region includes London, West Midlands, East of England, South East, East Midlands. More information about Reginal ethnical diversity can be access via https://www.ethnicity-facts-figures.service.gov.uk/uk-population-by-ethnicity/national-and-regional-populations/regional-ethnic-diversity/latest.

3.2 Model Specification

To test the relation between workforce DEI and patient safety incident, we use panel data regression analysis with trust and year fixed effect model. The baseline model is as follows:

$$\textit{Patient Safety Incidents}_{i,t+1} = \beta_1 \textit{ Workforce DEI}_{i,t} + \sum \beta_m \textit{ Controls}_{i,t} + \sum \beta_i \textit{ Trust}_i + \sum \beta_q \textit{ Year}_q + e_{it} \qquad \text{(Eq. 1)}$$

Where i denotes HNS trust and t denotes year. As discussed above, we use *Patient Safety Incidents*, defined as the number of patient safety incidents reported per 1000 admissions. The coefficient of workforce DEI, β_1 is our primary interest. For the control variables, we include NHS trust level variables that have been documented to affect NHS trust performance and efficiency (Veronesi et al., 2023). We control the size of the hospital by *patient admissions*, the overall performance of the trust by *SHMI*, the healthiness of financial resources by private *fundings*, the complexity of management by the *size of staff,* and the busyness of the trust by *bed occupancy*. All right-hand side variables have been lagged to account for the time delay in the effects of new policies and initiatives within the NHS units.

Using the Fixed Effects (FE) model for panel data analysis provides a significant advantage over Ordinary Least Squares (OLS), primarily due to its ability to account for unobserved, time-invariant variables. Such variables, encompassing inherent entity characteristics that remain constant over time, are typically unmeasurable or unavailable in datasets. The FE model, by harnessing within-entity variations and removing the influence of time-invariant variables, reduces this omitted variable bias, thereby producing more reliable and accurate estimates. Hence, in scenarios with potential unobserved heterogeneity, the FE model is a superior estimation method. All regressions include trust and year-fixed effects, with NHS trust clustered standard errors.

To examine under what circumstances the relationship between workforce DEI and patient safety incidents holds. As discussed in the hypothesis development section, we set up four moderators to test the hypotheses: DEI information transparency, and organizational complexity, and regional patient race diversity. We expand our regression model as below:

$$\begin{aligned} \textit{Patient Safety Incidents}_{i,t+1} = {} & \beta_1 \textit{ Workforce DEI}_{i,t} + \beta_2 \textit{ Moderator}_{i,t} \\ & + \beta_3 \textit{ Workforce DEI}_{i,t} \times \textit{Moderator}_{i,t} + \sum \beta_m \textit{ Controls}_{i,t} \\ & + \sum \beta_i \textit{ Trust}_i + \sum \beta_q \textit{ Year}_q + e_{it} \end{aligned} \qquad \text{(Eq. 2)}$$

The coefficient of the interaction term, *workforce DEI* × *Moderator*, β_3, indicates how these moderators strengthen or weaken the effects of *workforce DEI* on *Patient safety incidents.*

Table 2 provides descriptive statistics for variables used in the analysis. Patient safety incidents show an average rate of 187.48 per 1,000 admissions, indicating a substantial variance with a standard deviation of 63.47. The wide range between the minimum (74.64) and the maximum (471.36) underscores differing safety standards or reporting practices across NHS trusts. Staff diversity exhibits a noteworthy pattern. On average, non-white clinical staff account for 10.98% of the workforce, while non-white non-clinical staff constitute 15.11%. The higher proportion of non-white individuals in non-clinical roles suggests the potential importance of workforce DEI. Interestingly, our gender pay gap analysis (female salary median – male salary median) indicates a median salary difference of -12.46%, implying that men, on average, earn more than women within these trusts. Online employee reviews indicate average experience score of 0.42. However, with a standard deviation of 1.17, this suggests a spectrum of employee experiences and potential areas for improvement. Patient care data reveals an average of 68.58 admissions per 1,000 patients. The SHMI mortality rate measure presents an average score of 2.03, approximating the national average. Financial analysis reveals an average natural logarithm of private funding at 3.70, but a high standard deviation of 6.27. As for resources, the average natural logarithm of staff size is 8.76, and the bed occupancy rate averages at 87%, hinting at high demand or limited capacity.

Table 2. Summary Statistics

Variable	Definition	Obs.	Mean	SD	P25	P75
Patient Safety Incidents	Number of patient safety incidents reported per 1000 admissions	600	187.48	63.47	147.63	213.74
Race Diversity Clinical_1 (%)	The ratio of non-white clinical staff at Senior level	600	10.98	9.84	4.10	15.30
Race Diversity Clinical_2 (%)	The ratio of non-white clinical staff at Middle level	600	21.69	15.56	9.32	30.73
Race Diversity Non-Clinical (%)	The ratio of non-white non-clinical staff at Support level	600	15.11	16.95	3.11	18.29
Gender Pay Gap (%)	The median difference in salaries between female and male employees	600	-12.46	6.88	-17.45	-7.60
DEI Experiences	Textual analysis results from online employee reviews	600	0.42	1.17	0	0.32
Admissions	Patient admissions per 1,000 patients	600	68.58	32.53	46.24	85.67
SHMI	A metric used to assess the mortality rate in NHS trusts in England. It is calculated by comparing the actual number of patients who die following hospitalization at a specific NHS trust with the number of deaths that would be expected based on average figures for England, given the characteristics of the patients treated there. A score of 1 indicates high mortality, 2 is average, and 3 signifies lower-than-expected mortality	600	2.03	0.46	2.00	2.00
Fundings	Natural log of private fundings	600	3.70	6.27	0.00	11.36
Size of Staff	Natural log of staff number of NHS trust	600	8.76	0.49	8.38	9.07
Bed Occupancy	Percentage of bed occupancy of NHS trust	600	0.87	0.06	0.83	0.91

3.3 Double Machine Learning Approach for Prediction and Inference

In this study, we concentrate on deriving reliable estimations and insights for the workforce DEI index, asserting its causal interpretation. Our main empirical investigation employs the Fixed Effects model with lagged independent variables, chosen for its effectiveness in managing time-invariant NHS trust-level characteristics, reducing the endogeneity concern tied to omitted variable bias and reverse causality. For testing the robustness of our findings, we resort to the DML approach.

We consider using DML to solve the covariate selection problem in our high-dimensional model:

$$y = d\gamma + X\beta + \varepsilon$$

Where y denotes Patient Safety Incidents;

d denotes workforce DEI index

X denotes vector of control variables that might need to be included.

We estimate the effect of the *workforce DEI* index on *Patient Safety Incidents* and construct a confidence interval for the size of this effect. Given the numerous controls that could be selected in X, including all of them in the model with only 600 observations could lead to unreliable estimates of the workforce DEI coefficient (γ). DML leverages machine learning methods for covariate selection, mitigating bias arising from non-data-based selection of *X*, and provides reliable inference for γ, the primary focus of our study.

The DML method, also known as cross-fit partialling out (PO), is particularly advantageous when dealing with a myriad of controls that affect both the cause and the outcome. It yields a root n-consistent estimator along with doubly robust inferential statistics (Belloni et al., 2014a, 2014b; Chernozhukov et al., 2018). Importantly, even in the presence of multiple causal variables, DML remains effective by allowing for the selection of distinct control variables for each causal variable and outcome (utilizing LASSOs method), thereby facilitating the derivation of robust inferential parameters.

DML is used in our study to estimate causal effects with enhanced accuracy and robustness. The process involves two main steps: cross-fitting and double machine learning. In cross-fitting, the data is divided into multiple folds, and predictive models are trained and evaluated on each fold separately, reducing dependency on a single training-test split and stabilizing the estimation process. Then, in the double machine learning step, two machine learning models are utilized: one for predicting the treatment assignment and the other for predicting the outcome variable. By adjusting for the predicted treatment and outcome, this approach effectively controls for confounding variables, resulting in more reliable estimates of the average treatment effects (ATEs). The combination of cross-fitting and double machine learning improves the validity of the causal effect estimates and provides researchers with robust results for decision-making. In this study, the full sample is randomly split into 10 subsamples of approximately equal size (k = 10). The DML process is reported in Appendix E.

4 Results

4.1 How Workforce DEI Affect Patient Safety Incidents

In our primary analysis, we investigate the association between workforce DEI and patient safety incidents. The baseline regression results using fixed effect panel data regression are presented in Table 3. Due to limited observations (600) and the risk of introducing bias with numerous controls, we employ a LASSO approach discussed in 3.1.2. Race Diversity in Clinical staff at senior levels, Race Diversity in Clinical staff at middle levels, and Race Diversity in Non-clinical staff at support levels are selected by LASSO.

Table 3 DEI Predictor Construction and the Relationship between Workforce DEI and Patient Safety Incidents

Variable	Incidents $_t$ (1)	Incidents $_t$ (2)	Incidents $_t$ (3)	Incidents $_t$ (4)	Incidents $_{t+1}$ (5)
	Post-OLS estimation	Post-OLS estimation	OLS	OLS	OLS
Race Diversity Clinical_1 (%)	-2.692*** (0.824)				
Race Diversity Clinical_2 (%)	-1.003 (0.732)				
Race Diversity Non-Clinical (%)	-1.519** (0.707)				
Gender Pay Gap (%)		-0.958** (0.424)			
DEI Experiences			-2.096** (1.062)		
Workforce DEI				-32.106*** (7.310)	-17.647** (8.756)
Admissions	-1.582*** (0.240)	-1.452*** (0.242)	-1.414*** (0.245)	-1.520*** (0.241)	-0.361 (0.303)
SHMI	-10.461**	-10.697**	-11.952**	-11.307**	0.988

	(4.903)	(4.987)	(4.897)	(4.796)	(5.943)
Fundings	-0.372 (0.373)	-0.378 (0.379)	-0.349 (0.374)	-0.277 (0.365)	0.213 (0.455)
Size of Staff	95.765*** (28.788)	85.794*** (29.168)	78.239*** (28.544)	83.740*** (27.874)	69.530* (38.336)
Bed Occupancy	7.343 (39.519)	6.988 (40.152)	13.340 (39.951)	7.807 (39.172)	12.350 (47.975)
Trust fixed effects	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes
Adj R^2	0.219	0.191	0.147	0.227	0.060
Obs.	600	600	600	600	472

Column (1) of Table 3 shows the post-OLS results. A one percent increase in the ratio of non-white clinical staff at the senior level is associated with a reduction of approximately 2.692 patient safety incidents per 1000 admissions, significant at the 0.01 level. However, the coefficient for Race Diversity in Clinical staff at middle levels, selected by LASSO as -1.003, is not statistically significant in the post-OLS estimation, yet it maintains an inverse relationship with patient safety incidents. For race diversity in non-clinical staff, our result reveals that a one percent increase in the ratio of non-white non-clinical staff at the support level is associated with a reduction of approximately 1.519 patient safety incidents per 1000 admissions, significant at the 0.05 level. These findings suggest that while the impact of racial diversity varies across different staff levels and roles within NHS, increasing diversity, especially among non-white non-clinical support staff, consistently correlates with improved patient safety outcomes. This underscores the importance of implementing targeted diversity policies not only among clinical staff but also within support roles to enhance overall patient care and safety.

In Column (2), we present the results for gender DEI. Our analysis includes data on the gender pay gap and diversity ratios. Notably, the NHS workforce is composed of approximately 77% women, with 45% of clinical staff being women. This closely aligns with the national working population's female representation of 47%, as reported in the Gender in the NHS infographic (2019)[3]. Recently, the focus on gender equality within the NHS has been

[3] https://www.nhsemployers.org/articles/gender-nhs-infographic

predominantly on addressing the gender pay gap. Our findings show that diversity ratios vary insignificantly across different staff levels. The gender pay gap, defined as the median salary difference between male and female employees, is inversely correlated with the occurrence of patient safety incidents. Intriguingly, our data suggest that a 1% reduction in the gender pay gap can lead to a decrease of 0.958 patient safety incidents per 1000 admissions. This emphasizes the importance of achieving gender pay parity in contributing to enhanced patient safety within the NHS.

Column (3) analyzes the impact of DEI experiences within the workforce on patient safety incidents. The significant negative coefficient suggests that positive DEI experiences contribute to a reduction in patient safety incidents. This underscores the effectiveness of using a word embedding model to gain insights into the internal work environment. Essentially, improved DEI experiences are inversely related to the number of patient safety incidents. Specifically, an increase of one standard deviation in "Experiences" correlates with a decrease of approximately 2.452 patient safety incidents per 1000 admissions. This result is calculated by multiplying the coefficient of 2.096 by the standard deviation of 1.17. This highlights the importance of a positive, equitable, and inclusive work environment in enhancing patient safety and well-being within healthcare settings.

Column (4) presents the impact of the workforce DEI Index on patient safety incidents. This index represents the first principal component derived from a PCA analysis of several variables: *Race Diversity in Clinical roles, Race Diversity in Non-clinical roles, the Gender Pay Gap*, and *workforce DEI Experiences*. The *workforce DEI Index* is then used as an independent variable in a regression model predicting patient safety incidents, with other controls, trust and year-fixed effects. The coefficient associated with the workforce DEI Index is negative and statistically significant using full sample (600 observations). This suggests that an enhancement in the workforce DEI correlates with a reduction in the number of patient safety incidents. Essentially, a more diverse, equitable, and inclusive workforce appears to contribute positively towards patient safety within the healthcare setting.

To further address concerns of endogeneity, specifically reverse causality, and to account for the lagged effects of new initiatives, we re-estimated the model using the lagged values of the DEI index. By doing so, we lose one year of observation, reducing our sample size to 472. The results presented in Column (5) indicate a negative relationship between workforce DEI and patient safety incidents.

4.2 Predicting Patient Safety Incident Using Double Machine Learning

We are interested in how the model performs in predicting out-of-sample observations. Our study employs the DML approach, which enhances the precision and robustness of causal effect estimation. This technique comprises two crucial stages: cross-fitting and the application of two separate machine learning models.

During cross-fitting, we divide the dataset into multiple segments, or “folds.” This process mitigates reliance on a single train-test split by providing stability in the estimation process through individual training and evaluation on each fold. In the DML phase, one machine learning model forecasts treatment assignments while another predicts the outcome variable. Correcting for the predicted treatment and outcome variables helps manage confounding factors and yields reliable ATE estimates. The combination of cross-fitting and double machine learning not only augments the credibility of the causal effect estimates but also produces robust results.

Table 4 presents the results of our DML methodology applied to predict patient safety outcomes using workforce DEI as the primary predictor (Column 1). To explore potential nonlinear effects, we introduced the squared term of workforce DEI in Column (2). Our analysis involved dividing the dataset into 10 nearly equal-sized subsamples (corresponding to 10-fold cross-validation) to validate our model. Our model specification incorporates 148 control variables, including factors such as the trust location's racial diversity, trust incident rates, DEI policies, types of trusts, information environment, size, as well as trust and year-fixed effects. Of these, DML selected 20 controls for inclusion in Column (1) and 30 controls in Column (2). Our findings suggest that a unit score increase in the workforce DEI is associated with a decrease of 8.108 patient safety incidents per 1000 admissions. The 95% confidence interval for this estimate ranges from -13.167 to -3.070, indicating statistical significance. However, the squared term of workforce DEI did not show a significant influence, which may be attributed to the limited observations in our sample.

Table 4. Predicting Patient Safety Using Workforce DEI: Double Machine Learning

Variable	Incidents $_{t+1}$ (1)	Incidents $_{t+1}$ (2)
	DML	DML
Workforce DEI	-8.108^{***} (2.571) [-13.167, -3.070]	-6.992^{***} (2.477) [-11.847, -2.137]
Workforce DEI squared		-0.512 (1.070) [-2.609, 1.585]
Trust fixed effects	Yes	Yes
Year fixed effects	Yes	Yes
Covariates	Yes	Yes
Number of selected controls	20	30
Number of controls	148	148
Number of folds in cross-fit	10	10
Wald chi2(2)	9.95	8.43
Prob > chi2	0.0016	0.0118

4.3 Boundary Conditions for the Workforce DEI-Patient Safety Incident Link

we report the results examining whether and how boundary conditions moderate the negative relationship between workforce DEI and patient safety incidents. Table 5 presents the regression results for the possible boundary conditions as moderators in Columns (1)-(4). Model 1 reveals that the significant negative coefficient for the interaction term indicates that the effect of workforce DEI on patient safety incidents is influenced by regional race diversity.

Specifically, in regions with high racial diversity, the negative relationship between workforce DEI and patient safety incidents is even more pronounced. In contrast, the interaction term for workforce DEI × DEI information transparency in Column (2) is not statistically significant, suggesting that the effect of workforce DEI on patient safety incidents does not differ significantly with the level of information transparency[4]. Furthermore, the positive coefficient for the interaction term in Column (3) indicates that within complex NHS trusts, the negative relationship between workforce DEI and patient safety incidents is further mitigated. To ensure robustness, we defined Organizational Complexity as cases where the number of hospital sites and staff numbers exceed the median, respectively. The untabulated results align closely with those using NHS's self-reported categories.

4.4 Additional Analysis on DEI Policy Pressure

We consider DEI policy pressure as a boundary condition at the national level. Beginning in 2019, the UK government initiated a significant DEI policy shift with the publication of the Diversity and Inclusion Strategy 2019 to 2023 in April 2019. In response to this policy (see Appendix G), the NHS People Plan in 2020 embarked on developing crucial strategies and actionable plans to advocate for DEI in the NHS workforce. To quantitatively capture the influence of this external policy pressure, we integrated a dichotomous variable into our model, denoted as DEI policy pressure. This variable is set to one from 2020 onward, symbolizing the

---

[4] In this analysis, we initially assumed a linear moderating effect of transparency on the relationship between workforce DEI and patient safety incidents. However, as discussed in our hypothesis development, the impact of transparency might be non-linear, potentially leading to negative backlash when increased to maximum levels. Following Schilke (2014), we constructed an additional empirical model to test this non-linear moderating effect. We categorized the transparency variable into three levels: 0 for no disclosure, 1 for partial disclosure (for periods less than 3 years), and 2 for full disclosure. The results detailed in Table F-1 (see Appendix F) suggest that the relationship between workforce DEI and patient safety incidents varies across different levels of transparency in a quadratic manner. However, it is important to acknowledge that NHS trust disclosures often resemble a box-ticking process, predominantly driven by compliance rather than the provision of insightful information. This often results in disclosures that, while meeting regulatory requirements, offer limited interpretive value and pose significant information acquisition costs for external parties. Due to the lack of a continuous transparency variable, we are unable to visually demonstrate these effects, and thus the results should be interpreted with caution.

activation of these strategic objectives. This methodological decision enables us to empirically assess the potential ramifications of these policy changes in our analyses and results. Our results reveal that the significant negative coefficient for the interaction term (workforce DEI × DEI policy pressure) in Column (4) implies that the effect of workforce DEI on patient safety incidents is stronger (more negative) when policy pressure is present. This finding indicates that national policy plays a crucial role in amplifying the impact of DEI on patient safety. For policymakers, it demonstrates the significant change that can be accomplished through strategic policy shifts. For healthcare organizations, it underscores the need to be proactive and responsive to national DEI directives and initiatives.

Table 5. Workforce DEI and Patient Safety: Moderation effects

Variable DV	F.incidents_all (1)	F.incidents_all (2)	F.incidents_all (3)	F.incidents_all (4)
Workforce DEI	11.634 (13.170)	-18.088* (10.018)	-35.047*** (10.518)	-19.008** (7.634)
Workforce DEI × Regional patient race diversity	-45.574*** (15.448)			
DEI Information Transparency		-37.988* (21.594)		
Workforce DEI × DEI information transparency		0.033 (9.080)		
Workforce DEI × Organizational complexity			33.640*** (11.540)	
DEI policy pressure				44.691*** (6.970)
Workforce DEI × DEI policy pressure				-7.935*** (1.638)
Admissions	-0.373 (0.300)	-0.366 (0.303)	-0.238 (0.303)	-1.669*** (0.237)
SHMI	0.646 (5.877)	0.718 (5.935)	-1.518 (5.940)	-12.151*** (4.687)
Fundings	0.145 (0.451)	0.204 (0.455)	0.203 (0.450)	-0.302 (0.357)
Size of Staff	75.568** (37.961)	69.822* (38.298)	59.107 (38.085)	94.921*** (27.319)
Bed Occupancy	1.725 (47.572)	13.223 (48.262)	9.004 (47.464)	18.483 (38.318)
Trust fixed effects	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes
R-squared	0.068	0.060	0.072	0.070
Obs.	472	472	472	472

5 Discussion and Implication

5.1 Theoretical Contributions

The emphasis on DEI in the workplace has evolved beyond mere moral and social imperatives. Our findings illustrate that DEI directly impacts patient safety outcomes, a critical metric for healthcare providers. As the healthcare landscape continues to evolve, it is essential for healthcare institutions and policymakers to understand and leverage the potential of DEI in enhancing patient outcomes. This paper makes significant contributions to healthcare analytics by employing double machine learning and deep learning-enabled analytics to demonstrate the effectiveness of DEI as a proactive strategy for reducing patient safety incidents.

First, although the benefits of a diverse healthcare workforce are widely acknowledged, recent empirical research underscores persistent challenges (Hammond et al., 2022). A significant practical challenge in promoting DEI is the systematic understanding of the risks and opportunities associated with equity in the healthcare workforce (Rotenstein et al., 2021). This challenge may impede the sensemaking of DEI issues, which policymakers have identified as a primary task for implementing DEI practices in healthcare systems (Grubbs, 2020). To offer potential solutions, we synthesize data from various sources, including official workforce datasets, staff annual surveys, and data from third-party social media sites such as Glassdoor and Indeed. This data helps construct demographic and experiential measures of workforce DEI, serving as monitoring metrics to guide healthcare practitioners in establishing best DEI practices. Importantly, we introduce a novel approach to measuring DEI experience as a core component of workforce DEI, leveraging text analytics and NLP to analyze large-scale unstructured employee review content. This approach provides an alternative to traditional survey-based DEI experience measures (e.g., Simsekler and Qazi, 2022), which often suffer from social-desirability bias, sampling errors, and high costs. Not only does this method demonstrate the value of advanced analytics in uncovering patterns and insights within the healthcare sector, but it also enables researchers to capture the "actual voice" of healthcare professionals from social media platforms. In doing so, it responds to Ram and Goes's (2021) call for programmatic research in IS that employs cutting-edge analytical methods to achieve a deeper understanding of the healthcare sector.

Second, amid ongoing debate on the potential benefits and drawbacks of advocating DEI in the workplace (Romansky et al., 2021), it is crucial to examine the effects of workforce DEI on reducing patient safety incidents. This inquiry builds on previous research that empirically tests the board-level effects of diversity or equity on organizational performance (Bernile et al., 2018). However, establishing causal inference regarding how workforce DEI impacts patient safety is challenging due to our limited sample size and potential omitted variable biases. Initially, we used panel data analysis with year-fixed effects to mitigate omitted variable biases by controlling for time-invariant variables. Nonetheless, there remains a concern regarding uncontrolled time-variant variables that may concurrently influence patient safety and workforce DEI, potentially skewing the outcomes. To enhance robustness and address this endogeneity, we employed DML. DML excels in causal inference by adeptly managing high-dimensional covariates and confounders. Through machine learning, it effectively tackles nuisance functions and emphasizes orthogonality, ensuring that even if a model aspect is misspecified, our estimates remain consistent, a feature termed "double robustness." Additionally, DML minimizes the potential for overfitting through cross-fitting, making it versatile and suitable for various experimental designs. By applying this cutting-edge analytical approach to examine the impact of DEI on patient safety outcomes, our study makes an important contribution to advancing health analytics research (Baird et al., 2023; Marabelli and Chan, 2024).

Third, this research extends the DEI literature by emphasizing the pivotal role of contextual factors, a concept alluded to by Tang (2024) but not extensively explored previously. We found that the relationship between workforce DEI and patient safety incidents is more pronounced in regions characterized by high racial diversity. This finding challenges the prevalent “one-size-fits-all” paradigm and suggests that the effectiveness of DEI initiatives depends on the demographic composition of the local population. Such insights highlight the need for tailored DEI strategies, particularly in racially diverse regions where interventions may yield more significant benefits. Additionally, our research reveals a critical nuance in how organizational complexity affects DEI outcomes. Specifically, in more complex NHS trusts, the effectiveness of workforce DEI initiatives in reducing patient safety incidents is diminished compared to less complex organizations. This contribution to the DEI literature illustrates that increasing organizational complexity can weaken the positive impact of DEI initiatives on patient safety outcomes. Interestingly, our study also presents a counterintuitive insight regarding the interaction between DEI and information transparency. Despite previous emphasis on the importance of transparency in IS research (Granados & Gupta, 2013), the observed non-significant differential effect suggests a potential equilibrium. This implies that while information transparency remains critical, its presence might not enhance DEI’s impact on patient safety, especially when other DEI components are robustly integrated within the organization. These insights underline the importance of a holistic approach to DEI, considering both contextual and organizational factors to maximize its benefits.

5.2 Practical implications

Our study reveals significant implications for healthcare organizations, particularly within NHS institutions, as well as for healthcare policymakers. ***First***, our findings demonstrate a significant reduction in patient safety incidents associated with increased racial diversity, especially in non-clinical BME support roles and among senior clinical BME staff. This result aligns with the NHS’s 2023 Equality, Diversity, and Inclusion Improvement Plan (NHS England, 2023), which highlights that greater leadership diversity correlates with improved financial performance and enhanced patient satisfaction. However, our findings extend this perspective by quantitatively illustrating the safety benefits of diversity across various staffing levels, not just at the senior leadership level. Based on these insights, NHS institutions could expand their recruitment strategies and inclusivity training programs to focus more on entry-level positions, ensuring that diversity permeates all organizational levels. Encouraging collaboration across different roles and departments can also foster a more integrated approach to DEI, recognizing the intersectionality of staff identities and experiences. Joint initiatives between clinical and non-clinical staff to address specific DEI goals may further enhance both organizational cohesion and patient safety outcomes.

Second, the 2023 NHS Equality, Diversity, and Inclusion Improvement Plan recognizes the gender pay gap, particularly among medical staff and senior leaders, and outlines specific steps to mitigate these disparities (NHS England, 2023). While the plan typically frames the gender pay gap within the contexts of equity and regulation, our research introduces an important dimension by directly linking the gender pay gap to patient safety outcomes. This linkage implies that reducing the gender pay gap could lead not only to greater equity but also to notable improvements in clinical outcomes. Consequently, these findings could motivate the NHS to intensify efforts to address pay disparities, treating them as critical components of patient safety strategies rather than merely issues of compliance or ethics. This perspective encourages policymakers to view gender pay gap initiatives as essential for enhancing healthcare quality and advocates for a comprehensive approach in which pay equity is fundamental to achieving superior clinical performance and patient care.

Third, the 2023 NHS Equality, Diversity, and Inclusion Improvement Plan, particularly through High Impact Action 6, is dedicated to fostering a discrimination-free workplace by proactively preventing bullying and violence (NHS England, 2023). Our research enhances this initiative by emphasizing the experiential aspects of DEI, focusing on the actual experiences and environments as reported by healthcare professionals. Our analysis of employee feedback (see Table B-1 in Appendix B) reveals that terms such as “diversity,” “improvement,” “autonomy,” “prospects,” “organised,” and “supported” frequently appear. This finding underscores the necessity of prioritizing empowerment, recognition, and career advancement within DEI efforts. These findings should prompt NHS institutions and policymakers to broaden their DEI strategies, integrating comprehensive support systems that ensure staff feel valued and have access to clear, equitable career progression opportunities.

Finally, the strong relationship between workforce DEI and patient safety incidents, particularly in regions with high racial diversity, underscores the necessity for healthcare organizations in these areas to prioritize DEI efforts. This finding highlights the need for region-specific policies and frameworks to ensure that healthcare services are adequately equipped to serve diverse populations effectively. Additionally, in healthcare organizations with higher levels of organizational complexity, the effectiveness of workforce DEI initiatives in reducing patient safety incidents is diminished compared to organizations with lower complexity, points to the challenges of implementing DEI strategies in complex environments. Larger, multifaceted organizations may face inherent barriers that can interfere with the successful integration and impact of DEI initiatives. Given these findings, it is advisable for the NHS to develop a DEI dashboard as part of its improvement plan. This dashboard would aggregate key DEI metrics by region and trust, allowing local organizations to monitor progress, identify challenges, and engage in peer-to-peer learning. Implementing such a tool would support the execution of region-specific DEI strategies and enable health administrators to effectively track the impact of DEI initiatives, ensuring alignment with the specific needs of each region.

5.3 Limitations and Future Research

Our investigation into the role of DEI in influencing patient safety outcomes, while insightful, is accompanied by some limitations. Addressing these limitations not only strengthens the validity of our findings but also opens promising avenues for future research. One primary limitation is the data sources utilized. Though we drew from official NHS workforce datasets and third-party platforms (e.g., Glassdoor and Indeed), each source might possess inherent biases. For instance, Glassdoor/Indeed data may not be fully representative, potentially omitting voices of employees not actively using these platforms. To offer a more holistic picture, future research could examine additional or alternative data sources, such as direct interviews or other employee review platforms, ensuring a wider coverage of DEI experiences.

While our study shed light on the gender pay gap, we potentially skirted around deeper-seated gender biases such as positional disparities or underrepresentation in leadership roles. This paves the way for future research to conduct a more granular examination of gender biases, going beyond pay disparities to identify and address broader gender-related challenges. In addition, our exploration into contextual factors focused on regional racial diversity and organizational complexity. Yet, the universe of contextual factors is vast, and others such as socioeconomic status or linguistic diversity remain uncharted in this study. Future studies could explore these uncharted areas to reveal how different contextual factors interplay with DEI initiatives and influence patient safety outcomes. Finally, our study is based on NHS data from England, which may limit the applicability of our findings to other healthcare systems or regions. Future research could explore similar questions in diverse international settings to assess the external validity of our results.

Acknowledgments
This work was supported by the Research, Knowledge Exchange, and Impact Stimulation Funding from the Behavioural Research for Inclusivity, Sustainability and Technological Transformation (BRISTT) and the Centre for Research into Accounting and Finance in Context (CRAFiC) at the Sheffield University Management School, University of Sheffield, UK.

References

Acquisti, A., & Fong, C. (2020). An experiment in hiring discrimination via online social networks. *Management Science*, 66(3), 1005-1024.

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, 94(2), 291-309.

Agarwal, R., Dugas, M., Gao, G., & Kannan, P. K. (2020). Emerging technologies and analytics for a new era of value-centered marketing in healthcare. *Journal of the Academy of Marketing Science*, 48, 9-23.

Anderson, P. (1999). Perspective: Complexity theory and organization science. *Organization Science*, 10(3), 216-232.

Andrevski, G., Richard, O. C., Shaw, J. D., & Ferrier, W. J. (2014). Racial diversity and firm performance: The mediating role of competitive intensity. *Journal of Management*, 40(3), 820-844.

Arsel, Z., Crockett, D., & Scott, M. L. (2022). Diversity, equity, and inclusion (DEI) in the Journal of Consumer Research: A curation and research Agenda. *Journal of Consumer Research*, 48(5), 920-933.

Baird, A., Xia, Y., & Kohli, R. (2023), Health Analytics and IS Theorizing. Retrieved from 21 Feburary https://aisnet.org/news/news.asp?id=624886

Bardhan, I., Chen, H., & Karahanna, E. (2020). Connecting systems, data, and people: A multidisciplinary research roadmap for chronic disease management. *MIS Quarterly*, 44(1), 185-200.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014a). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives*, 28(2), 29-50.

Belloni, A., Chernozhukov, V., & Hansen, C. (2014b). Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2), 608-650.

Bennedsen, M., Simintzi, E., Tsoutsoura, M., & Wolfenzon, D. (2022). Do firms respond to gender pay gap transparency?. *The Journal of Finance*, 77(4), 2051-2091.

Bernile, G., Bhagwat, V., & Yonker, S. (2018). Board diversity, firm risk, and corporate policies. *Journal of Financial Economics*, 127(3), 588-612.

Bird, S., Klein, E., & Loper, E. (2009). Natural language processing with Python: analyzing text with the natural language toolkit. O'Reilly Media, Inc.

Bishu, S. G., & Alkadry, M. G. (2017). A systematic review of the gender pay gap and factors that predict it. *Administration & Society*, 49(1), 65-104.

Boulware, L. E., Corbie, G., Aguilar-Gaxiola, S., Wilkins, C. H., Ruiz, R., Vitale, A., & Egede, L. E. (2022a). Combating structural inequities—Diversity, equity, and inclusion in clinical and translational research. *New England Journal of Medicine*, 386, 201-203.

Boulware, L., Vitale, A., Ruiz, R., Corbie, G., Aguilar-Gaxiola, S., Wilkins, C. H., & Egede, L. E. (2022b). Diversity, equity and inclusion actions from the NCATS Clinical and Translational Science awarded programs. *Nature Medicine*, 28(9), 1730-1731.

British Medical Association (2022). Racism in medicine. Available at: https://www.bma.org.uk/media/5746/bma-racism-in-medicine-survey-report-15-june-2022.pdf.

Campbell, D. W., & Shang, R. (2022). Tone at the bottom: Measuring corporate misconduct risk from the text of employee reviews. *Management Science*, 68(9), 7034-7053.

Chang, E. H., Kirgios, E. L., Rai, A., & Milkman, K. L. (2020). The isolated choice effect and its implications for gender diversity in organizations. *Management Science*, 66(6), 2752-2761.

Chattopadhyay, P., Tluchowska, M., & George, E. (2004). Identifying the ingroup: A closer look at the influence of demographic dissimilarity on employee social identity. *Academy of Management Review*, 29(2), 180-202.

Chen, J., & Chen, Z. (2008). Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*, *95*(3), 759-771.

Chen, L., Lee, H.L. & Tang, C.S. (2022) Supply chain fairness. *Production and Operations Management,* 31(2), 4304-4318.

Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1-C68.

Cumming, D., Leung, T. Y., & Rui, O. (2015). Gender diversity and securities fraud. *Academy of Management Journal*, 58(5), 1572-1593.

Dube, S., & Zhu, C. (2021). The disciplinary effect of social media: Evidence from firms' responses to Glassdoor reviews. *Journal of Accounting Research*, 59(5), 1783-1825.

Earley, C. P., & Mosakowski, E. (2000). Creating hybrid team cultures: An empirical test of transnational team functioning. *Academy of Management Journal*, 43(1), 26-49.

Ferner, A., Almond, P., & Colling, T. (2005). Institutional theory and the cross-national transfer of employment policy: The case of 'workforce diversity'in US multinationals. *Journal of International Business Studies*, *36*, 304-321.

Gillespie, A., & Reader, T. W. (2023). Online patient feedback as a safety valve: An automated language analysis of unnoticed and unresolved safety incidents. *Risk analysis*, *43*(7), 1463-1477.

Ginsburg, L., Castel, E., Tregunno, D., & Norton, P. G. (2012). The H-PEPSS: an instrument to measure health professionals' perceptions of patient safety competence at entry into practice. *BMJ Quality & Safety*, 21(8), 676-684.

Gottlieb, A. S., & Jagsi, R. (2021). Closing the gender pay gap in medicine. *New England Journal of Medicine*, 385(27), 2501-2504.

Granados, N., & Gupta, A. (2013). Transparency strategy: Competing with information in a digital world. *MIS Quarterly*, 37(2), 637-641.

Grubbs, V. (2020). Diversity, equity, and inclusion that matter. *New England Journal of Medicine*, 383(4), e25.

Guillaume, Y. R., Dawson, J. F., Otaye-Ebede, L., Woods, S. A., & West, M. A. (2017). Harnessing demographic differences in organizations: What moderates the effects of workplace diversity?. *Journal of Organizational Behavior*, 38(2), 276-303.

Hales, J., Moon Jr, J. R., & Swenson, L. A. (2018). A new era of voluntary disclosure? Empirical evidence on how employee postings on social media relate to future corporate disclosures. *Accounting, Organizations and Society*, 68, 88-108.

Hammond, J., Davies, N., Morrow, E., Ross, F., Vandrevala, T., & Harris, R. (2022). Raising the curtain on the equality theatre: a study of recruitment to first healthcare job post-qualification in the UK National Health Service. *Human Resources for Health*, 20(1), 1-16.

Hellerstedt, K., Uman, T., & Wennberg, K. (2024). Fooled by diversity? When diversity initiatives exacerbate rather than mitigate bias and inequality. *Academy of Management Perspectives*, *38*(1), 23-42.

Hoff, T. (2021). The gender pay gap in medicine: a systematic review. *Health Care Management Review*, 46(3), E37-E49.

Hogg, M.A., & Terry, D. J. (2000). Social identity and self-categorization processes in organizational context. *Academy of Management Review*, 25(1), 121-140.

Hu, X., & Casey, T. (2021). How and when organization identification promotes safety voice among healthcare professionals. *Journal of Advanced Nursing*, 77(9), 3733-3744.

Huang, K., Li, M., & Markov, S. (2020). What do employees know? Evidence from a social media platform. *The Accounting Review*, 95(2), 199-226.

Hydari, M. Z., Telang, R., & Marella, W. M. (2019). Saving patient Ryan—can advanced electronic medical records make patient care safer?. *Management Science*, 65(5), 2041-2059.

Jackson, S. R. (2023). (Not) Paying for Diversity: Repugnant Market Concerns Associated with Transactional Approaches to Diversity Recruitment. *Administrative Science Quarterly*, *68*(3), 824-866.

Joshi, A., Liao, H., & Roh, H. (2011). Bridging domains in workplace demography research: A review and reconceptualization. *Journal of Management*, 37(2), 521-552.

Kapadia, D., Zhang, J., Salway, S., Nazroo, J., Booth, A., Villarroel-Williams, N., ... & Esmail, A. (2022). Ethnic inequalities in healthcare: a rapid evidence review. NHS Race and Health Observatory. Available at: https://www.nhsrho.org/wp-content/uploads/2023/05/RHO-Rapid-Review-Final-Report_.pdf.

Lee, G., & Xia, W. (2010). Toward agile: an integrated analysis of quantitative and qualitative field data on software development agility. *MIS Quarterly*, 34(1), 87-114.

Li, F., Lo, C. K., Tang, C. S., & Zhou, P. (2024). Will Diversity, Equity, and Inclusion Commitment Improve Manufacturing Firms' Market Performance? A Signaling Theory Perspective on DEI

Announcements. *Production and Operations Management*. https://doi.org/10.1177/10591478241265479.

Li, K., Mai, F., Shen, R. & Yan, X. (2021). Measuring corporate culture using machine learning. *The Review of Financial Studies*, 34(7), 3265-3315.

Mael, F., & Ashforth, B. E. (1992). Alumni and their alma mater: A partial test of the reformulated model of organizational identification. *Journal of Organizational Behavior*, 13(2), 103-123.

Marabelli, M., & Chan, Y. E. (2024). The strategic value of DEI in the information systems discipline. *The Journal of Strategic Information Systems*, *33*(1), 101823.

Mas, A. (2016). Does Disclosure Affect CEO Pay Setting?: Evidence from the Passage of the 1934 Securities and Exchange Act. Princeton University, Industrial Relations Section.

Mas, A. (2017). Does transparency lead to pay compression?. *Journal of Political Economy*, 125(5), 1683-1721.

Molloy, A. (2022). Workforce race discrimination adversely impacts patient care too by Roger Kline. Available at: https://blogs.bmj.com/bmjleader/2022/05/31/workforce-race-discrimination-adversely-impacts-patient-care-too-by-roger-kline/.

Mullainathan, S. and Spiess, J., 2017. Machine learning: an applied econometric approach. *Journal of Economic Perspectives*, 31(2), pp.87-106

Nagesh, A. (2022). NHS racism making doctors 'anxious and depressed'. Available at: https://www.bbc.co.uk/news/uk-61792140.

Nancy, P., Currie, W., & Whitley, E. A. (2016). Entangled stakeholder roles and perceptions in health information systems: a longitudinal study of the UK NHS N3 network. *Journal of the Association for Information Systems*, *17*(2), 1.

NHS East of England (2021). No More Tick Boxes: A review of the evidence on how to make recruitment and career progression fairer. Available at: https://www.england.nhs.uk/east-of-england/wp-content/uploads/sites/47/2021/10/NHSE-Recruitment-Research-Document-FINAL-2.2.pdf.

NHS England (2022a). NHS workforce race equality standard: 2021 data analysis report for NHS trusts. Available at: https://www.england.nhs.uk/wp-content/uploads/2022/04/Workforce-Race-Equality-Standard-report-2021-.pdf.

NHS England (2022b). NRLS national patient safety incident reports: Commentary. Available at: https://www.england.nhs.uk/wp-content/uploads/2022/10/NAPSIR-commentary-Oct-22-FINAL-v4.pdf.

NHS England (2023). NHS equality, diversity, and inclusion improvement plan. Available at: https://www.england.nhs.uk/long-read/nhs-equality-diversity-and-inclusion-improvement-plan/.

Obloj, T., & Zenger, T. (2022). The influence of pay transparency on (gender) inequity, inequality and the performance basis of pay. *Nature Human Behaviour*, 6(5), 646-655.

Okuyama, A., Martowirono, K., & Bijnen, B. (2011). Assessing the patient safety competencies of healthcare professionals: a systematic review. *BMJ Quality & Safety*, 20(11), 991-1000.

Park, Y. W., Voss, G. B., & Voss, Z. G. (2023). Advancing customer diversity, equity, and inclusion: Measurement, stakeholder influence, and the role of marketing. *Journal of the Academy of Marketing Science*, 51, 174-197.

Patel, P. C., & Feng, C. (2021). LGBT workplace equality policy and customer satisfaction: The roles of marketing capability and demand instability. *Journal of Public Policy & Marketing*, 40(1), 7-26.

Ram, S., & Goes, P. (2021). Focusing on programmatic high impact information systems research, not theory, to address grand challenges. *MIS Quarterly*, 45(1), 479-483.

Randel, A. E., Dean, M. A., Ehrhart, K. H., Chung, B., & Shore, L. (2016). Leader inclusiveness, psychological diversity climate, and helping behaviors. *Journal of Managerial Psychology*, *31*(1), 216-234.

Ranganathan, C., Watson-Manheim, M. B., & Keeler, J. (2004). Bringing professionals on board: lessons on executing IT-enabled organizational transformation. *MIS Quarterly Executive*, 3(3), 151-160.

Robert Jr, L. P., Dennis, A. R., & Ahuja, M. K. (2018). Differences are different: Examining the effects of communication media on the impacts of racial and gender diversity in decision-making teams. *Information Systems Research*, 29(3), 525-545.

Romansky, L., Garrod, M., Brown, K., & Deo, K. (2021). How to measure inclusion in the workplace. *Harvard Business Review*, 27.

Rotenstein, L. S., Reede, J. Y., & Jena, A. B. (2021). Addressing workforce diversity—a quality-improvement framework. *New England Journal of Medicine*, 384(12), 1083-1086.

Sabharwal, M. (2014). Is diversity management sufficient? Organizational inclusion to further performance. *Public Personnel Management*, 43(2), 197-217.

Salton, G., & Buckley, C. (1988). Term-weighting approaches in automatic text retrieval. *Information Processing & Management*, *24*(5), 513-523.

Sautner, Z., Van Lent, L., Vilkov, G. and Zhang, R., 2023. Firm-level climate change exposure. *The Journal of Finance*, *78*(3), pp.1449-1498.

Schilke, O. (2014). On the contingent value of dynamic capabilities for competitive advantage: The nonlinear moderating effect of environmental dynamism. *Strategic Management Journal*, 35(2), 179-203.

Simsekler, M. C. E., & Qazi, A. (2022). Adoption of a data-driven Bayesian belief network investigating organizational factors that influence patient safety. *Risk Analysis*, 42(6), 1277-1293.

Stanford, F. C. (2020). The importance of diversity and inclusion in the healthcare workforce. *Journal of the National Medical Association*, 112(3), 247-249.

Tang, C. S. (2024). Diversity, equity, and inclusion: Decision science research opportunities. *Decision Sciences*, *55*(1), 7-16.

Tong, Y., Tan, C. H., Sia, C. L., Shi, Y., & Teo, H. H. (2022). Rural-urban healthcare access inequality challenge: Transformative roles of information technology. *MIS Quarterly*, 46(4), 1937-1982.

Tucker, A. L., Zheng, S., Gardner, J. W., & Bohn, R. E. (2020). When do workarounds help or hurt patient outcomes? The moderating role of operational failures. *Journal of Operations Management*, 66(1-2), 67-90.

Van Dijk, H., Van Engen, M. L., & Van Knippenberg, D. (2012). Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Organizational Behavior and Human Decision Processes*, 119(1), 38-53.

Van Knippenberg, D., De Dreu, C. K., & Homan, A. C. (2004). Work group diversity and group performance: an integrative model and research agenda. *Journal of Applied Psychology*, 89(6), 1008.

Veronesi, G., Kirkpatrick, I., Altanlar, A., & Sarto, F. (2023). Corporatization, administrative intensity, and the performance of public sector organizations. *Journal of Public Administration Research and Theory*, *33*(4), 701-715.

Wang, D., Zhu, Q., Avolio, B. J., Shen, W., & Waldman, D. (2022a). Do employees' views matter in corporate governance? The relationship between employee approval and CEO dismissal. *Strategic Management Journal*. https://doi.org/10.1002/smj.3465.

Wang, X., Yang, X., & Ye, J. (2022b). Mitigating IT professionals' turnover in non-IT organizations: an organizational identification perspective. *Journal of the Association for Information Systems*, 23(5), 1271-1295.

Xia, W., & Lee, G. (2005). Complexity of information systems development projects: conceptualization and measurement development. *Journal of Management Information Systems*, *22*(1), 45-83.

Xin, M., & Choudhary, V. (2019). IT investment under competition: The role of implementation failure. *Management Science*, 65(4), 1909-1925.

Zyphur, M. J., Voelkle, M. C., Tay, L., Allison, P. D., Preacher, K. J., Zhang, Z., ... & Diener, E. (2020). From data to causes II: Comparing approaches to panel data analysis. *Organizational Research Methods*, 23(4), 688-716.

Appendix A. A Comprehensive Listing of the EDI Variables Considered During the LASSO

Category	Variables	Source
Race demographic diversity for clinical staff	Support (Bands 1-4), Middle (Bands 5-7), Senior (Bands 8a-9), VSM (Very Senior Managers), by ethnicity (White and Black and Minority Ethnic; BME)	NHS-WRES
Racial differences in work experiences	*Recruitment and selection*: Relative likelihood of white applicants being appointed from shortlisting across all posts compared to BME applicants *Formal disciplinary:* Relative likelihood of BME staff entering the formal disciplinary process compared to white staff *Training and development*: Relative likelihood of white staff accessing non-mandatory training and continuous professional development (CPD) compared to BME staff *Equal opportunities*: Relative likelihood of BME staff believing that their trust provides equal opportunities for career progression or promotion compared to white staff	NHS-WRES
Race demographic diversity for nonclinical staff	Support (Bands 1-4), Middle (Bands 5-7), Senior (Bands 8a-9), VSM (Very Senior Managers), by ethnicity (White and BME)	NHS-WRES
Gender pay gap	*Gender pay gap in hourly pay:* Median % difference between male and female hourly pay (negative = women's mean hourly pay is higher) *Gender pay gap in bonus*: Median % difference between male and female bonus pay (negative = women's mean bonus pay is higher) Gender diversity at Lower Quartile, Lower Middle Quartile, upper middle quartile, and top Quartile	Reported in gender pay gap data in GPGS.

Note: NHS-WRES: Workforce Race Equality Standard; GPGS: Gender Pay Gap Service

Appendix B. Detailed Analysis of DEI-Related Term Frequency and Weighting in Employee Reviews

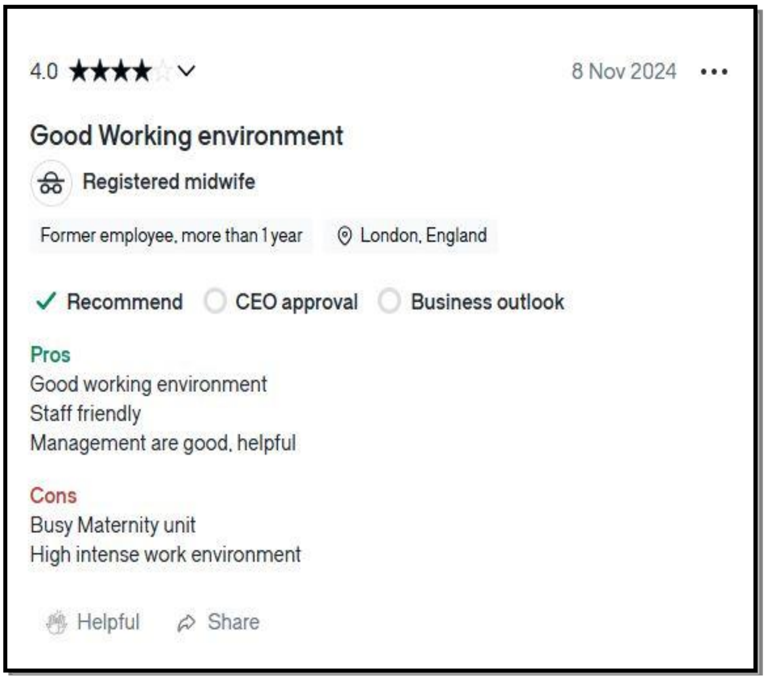


Figure B-1. Screenshot of an employee review in Glassdoor

Table B-1. Top 30 Term Frequency-Inverse Document Frequency (TF-IDF) Weighted DEI Words in Pros Section

Word	TF-IDF Score	Percent	Cumulative TF-IDF Score	Cumulative Percent
diversity	1.60	1.09	1.60	1.09
improvement	1.53	1.05	3.13	2.14
autonomy	1.50	1.03	4.63	3.17
towards	1.41	0.97	6.04	4.13
sense	1.40	0.96	7.45	5.09
train	1.38	0.94	8.83	6.04
valued	1.32	0.91	10.15	6.94
supported	1.32	0.90	11.47	7.85
organised	1.31	0.90	12.78	8.75
prospects	1.31	0.89	14.09	9.64
clinicians	1.26	0.86	15.35	10.50
including	1.22	0.83	16.56	11.33
driven	1.21	0.83	17.77	12.16
pharmacy	1.16	0.79	18.93	12.95
progressive	1.12	0.77	20.05	13.72
seniors	1.12	0.77	21.17	14.49
childcare	1.10	0.75	22.28	15.24
amongst	1.07	0.73	23.34	15.97
encourage	1.04	0.71	24.38	16.68
workforce	1.01	0.69	25.40	17.38
effective	1.01	0.69	26.41	18.07
fast	0.94	0.65	27.35	18.71

security	0.90	0.62	28.25	19.33
aspects	0.88	0.60	29.14	19.93
administration	0.88	0.60	30.01	20.53
internal	0.86	0.59	30.87	21.12
wages	0.86	0.59	31.73	21.71
innovation	0.83	0.57	32.56	22.27
environments	0.82	0.56	33.38	22.84
shop	0.82	0.56	34.19	23.39

Appendix C. The Results of the PCA Analysis

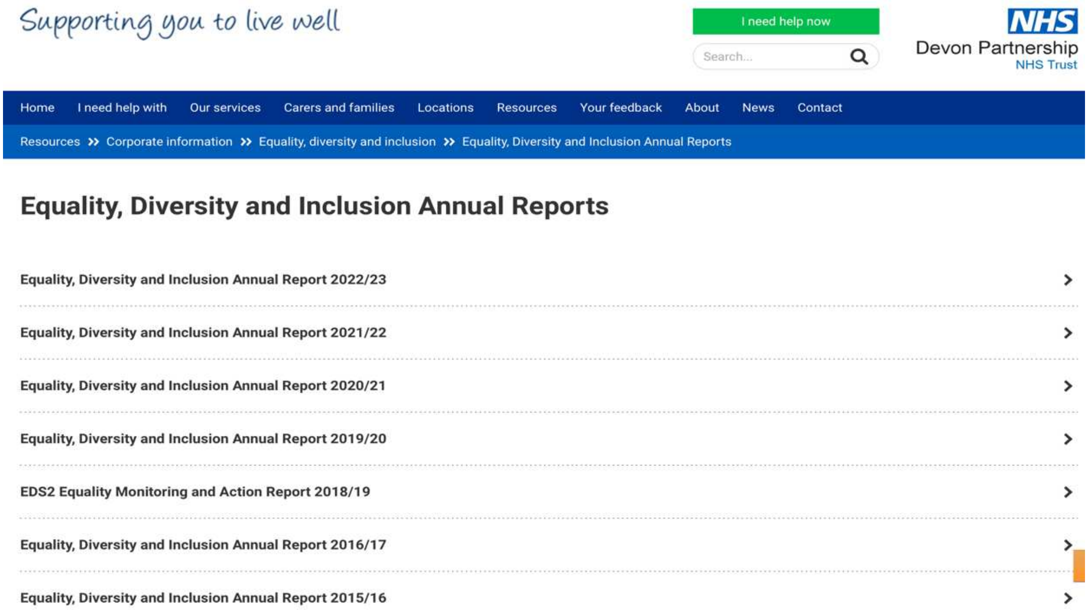
Panel A: Variable loading after rotation		
Variable	Factor1	Factor2
Race Diversity Clinical_1 (%)	0.571	-0.062
Race Diversity Clinical_2 (%)	0.584	-0.017
Race Diversity Non-Clinical (%)	0.564	0.103
Gender Pay Gap (%)	0.057	0.816
DEI Experience	-0.105	0.566
Panel B: Components derived from the PCA		
Factor	Eigenvalue	Proportion of Explained Variance
Factor1	2.742	0.548
Factor2	1.052	0.214
Factor3	0.949	0.190
Factor4	0.176	0.035
Factor5	0.082	0.016

Appendix D. The examples of full DEI information transparency and limited DEI information transparency

<table>
<tr>
<td>Full DEI information transparency (recorded as 1)</td>
<td>

Devon Partnership NHS Trust:

- A full list of equality, diversity, and inclusion annul reports between 2015 and 2023 is provided.



- A full list of workforce race quality standard (between 2018-2022) and gender pay gap reports (between 2017-2022) is provided.

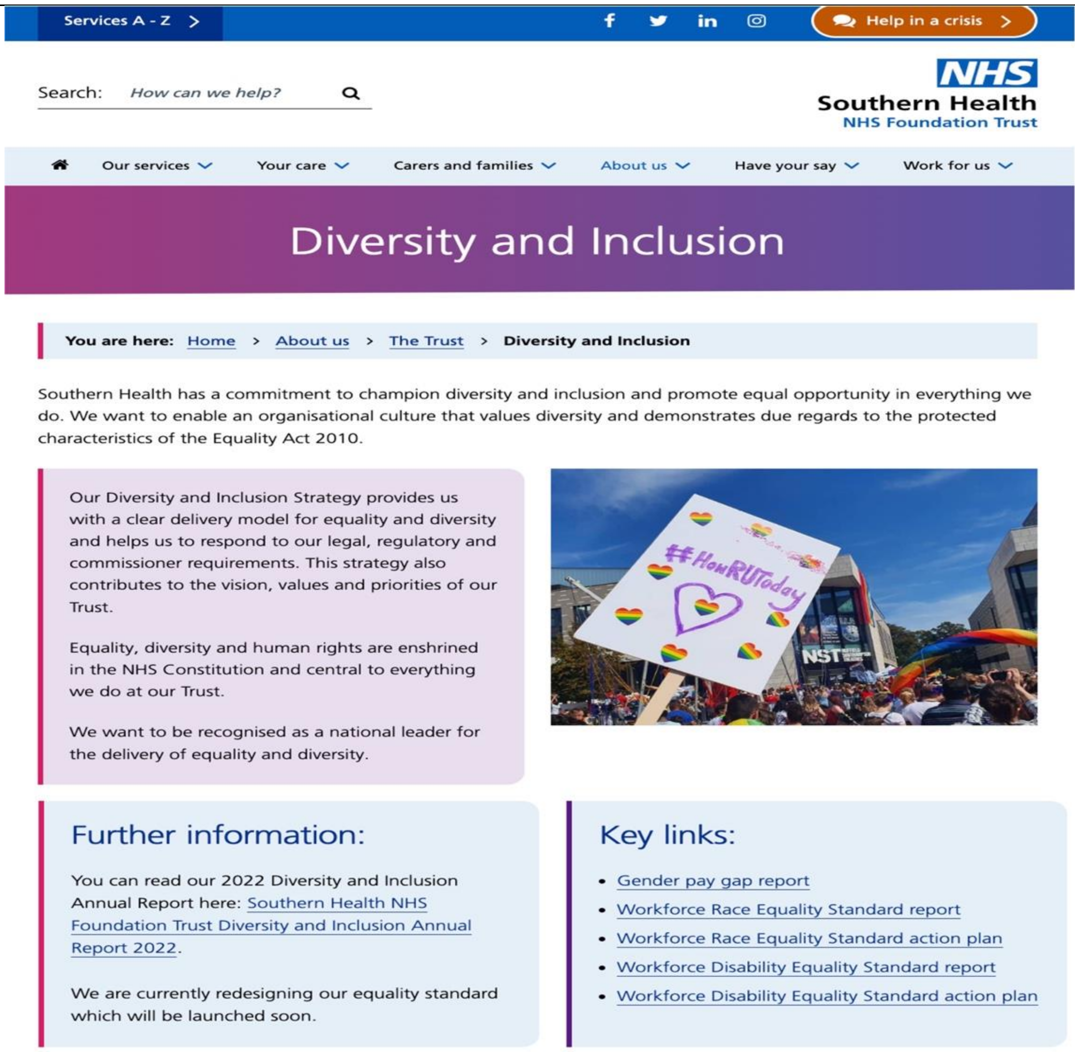
Source: https://www.dpt.nhs.uk/resources/corporate-information/equality-diversity-and-inclusion

</td>
</tr>
<tr>
<td>Limited DEI information transparency (recorded as 0)</td>
<td>

Southern Health NHS Foundation Trust:

- Restrict disclosure to data from only the most recent one or two years.

</td>
</tr>
</table>


Services A - Z >

Help in a crisis >

Search: How can we help?

NHS
Southern Health
NHS Foundation Trust

Our services | Your care | Carers and families | About us | Have your say | Work for us

Diversity and Inclusion

You are here: Home > About us > The Trust > Diversity and Inclusion

Southern Health has a commitment to champion diversity and inclusion and promote equal opportunity in everything we do. We want to enable an organisational culture that values diversity and demonstrates due regards to the protected characteristics of the Equality Act 2010.

Our Diversity and Inclusion Strategy provides us with a clear delivery model for equality and diversity and helps us to respond to our legal, regulatory and commissioner requirements. This strategy also contributes to the vision, values and priorities of our Trust.

Equality, diversity and human rights are enshrined in the NHS Constitution and central to everything we do at our Trust.

We want to be recognised as a national leader for the delivery of equality and diversity.

Further information:

You can read our 2022 Diversity and Inclusion Annual Report here: Southern Health NHS Foundation Trust Diversity and Inclusion Annual Report 2022.

We are currently redesigning our equality standard which will be launched soon.

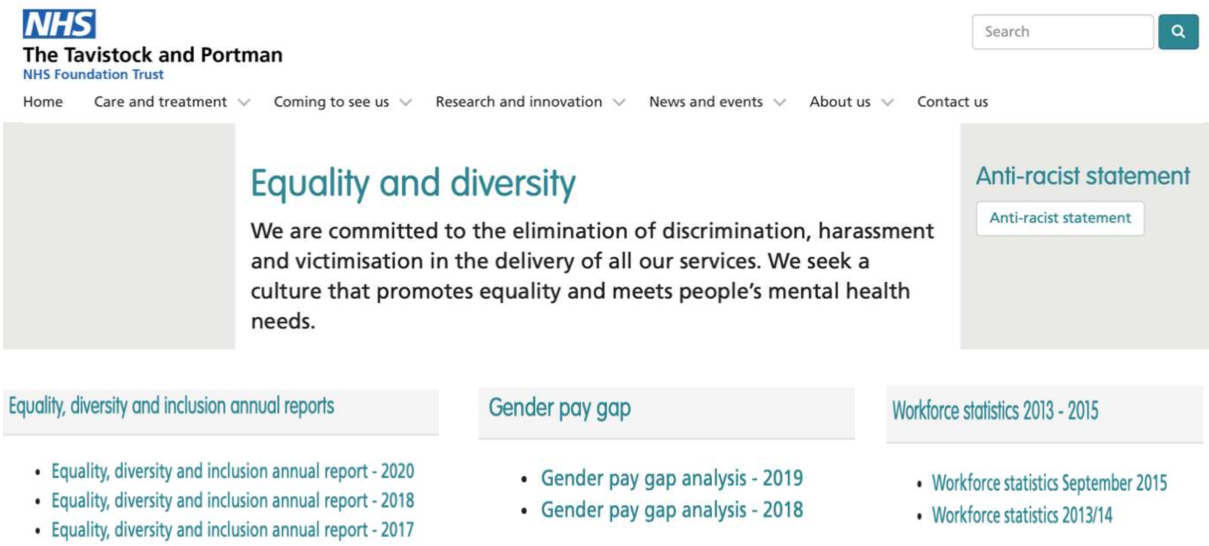
Key links:

- Gender pay gap report
- Workforce Race Equality Standard report
- Workforce Race Equality Standard action plan
- Workforce Disability Equality Standard report
- Workforce Disability Equality Standard action plan

Source: https://www.southernhealth.nhs.uk/about-us/trust/diversity-and-inclusion

The Tavistock and Portman NHS Foundation Trust:

- Several pieces of DEI information are either absent or outdated.


NHS
The Tavistock and Portman
NHS Foundation Trust

Search

Home | Care and treatment | Coming to see us | Research and innovation | News and events | About us | Contact us

Equality and diversity

We are committed to the elimination of discrimination, harassment and victimisation in the delivery of all our services. We seek a culture that promotes equality and meets people's mental health needs.

Anti-racist statement

Anti-racist statement

Equality, diversity and inclusion annual reports

- Equality, diversity and inclusion annual report - 2020
- Equality, diversity and inclusion annual report - 2018
- Equality, diversity and inclusion annual report - 2017

Gender pay gap

- Gender pay gap analysis - 2019
- Gender pay gap analysis - 2018

Workforce statistics 2013 - 2015

- Workforce statistics September 2015
- Workforce statistics 2013/14

Source: https://tavistockandportman.nhs.uk/about-us/who-we-are/equalities/

Appendix E. Double Machine Learning Process

DML, known as cross-fit partialling out, are deployed within k approximately equal-sized subsamples of the dataset to select controls for each of the causal variables as well as the outcome, in the model. The process can be described as follows.

DML process	Description
Step 1: Split the Data	Divide the dataset into multiple folds (e.g., 10 folds in this paper) to facilitate the cross-fitting procedure. This step helps avoid using the same data for both model training and validation.
Step 2: Model specification	For choosing the tuning parameters λ, which determines which covariates will be included and which will be excluded. The selection of tuning parameters (λ) plays a crucial role in determining which covariates will be included or excluded from the model. To achieve a balanced model with relevant covariates, we adopt the adaptive lasso approach. The adaptive lasso is a powerful technique chosen for its ability to effectively incorporate important covariates while managing model complexity. It strikes a balance between two key considerations: including relevant covariates and avoiding the inclusion of irrelevant ones. This capability allows the model to capture essential relationships, ensuring the estimation of accurate and reliable effects.
Step 3: Cross-fitting Loop:	For each fold in the dataset, divide the data into a training set and a validation set. Train the treatment effect model on the training set using only the instruments and covariates as features. Use the treatment effect model to predict the estimated treatment effects for the validation set. Train the outcome prediction model on the training set using the treatment effects and covariates as features. Use the outcome prediction model to predict the outcomes for the validation set.
Step 4: Orthogonalization	Calculate the residuals for both the treatment effect predictions and outcome predictions. Orthogonalization is a crucial step that reduces the risk of model selection bias and makes the estimates more robust.
Step 5: Estimation of Causal Effects	With the orthogonalized residuals, estimate the causal effect by fitting a simple linear regression model. The instrumental variable estimate of the causal effect is now obtained from this regression.
Step 6: Assessment and Inference	Assess the performance of your model using relevant metrics and perform statistical inference to determine the significance and confidence intervals of the estimated causal effects.
Step 7: Cross-Validation Averaging	Average the causal effect estimates obtained from the different folds to obtain the final causal effect estimate. This averaging helps reduce the variance and provides a more reliable estimate of the treatment effect.

Appendix F. Predictor Construction and the Relationship between Workforce DEI and Patient Safety Incidents: Non-linear Moderation of Transparency.

Following Schilke (2014), we construct an additional empirical model below to test the nonlinear moderating effect:

$$\begin{aligned} \textit{Patient Safety Incidents}_{i,t+1} = {} & \beta_1 \textit{ Workforce DEI}_{i,t} + \beta_2 \textit{ Transparency}_{i,t} \\ & + \beta_3 \textit{ Workforce DEI}_{i,t} \times \textit{Transparency}_{i,t} + \beta_4 \textit{ Transparency}^2_{i,t} \\ & + \beta_5 \textit{ Workforce DEI}_{i,t} \times \textit{Transparency}^2_{i,t} + \sum \beta_m \textit{ Controls}_{i,t} \\ & + \sum \beta_i \textit{ Trust}_i + \sum \beta_q \textit{ Year}_q + e_{it} \end{aligned}$$

The transparency variable is categorized into three levels: 0 represents no disclosure, 1 indicates partial disclosure for periods less than 3 years, and 2 signifies full disclosure. We are particularly focused on the coefficient β_5 . If β_5 is negative, it suggests that transparency exerts a moderating effect that follows a downward (concave) ∩ shape moderating effects, Conversely, a positive β_5 would indicate that transparency's moderating effect follows an upward (convex) U-shaped curve. Table G-1 shows the empirical results. The coefficient β_5 is negative and significant, indicating a concave ∩ shape moderating effects of transparency. It suggests that the relationship between workforce DEI and patient safety incident vary across difference levels of transparent in a quadratic manner. However, it is important to acknowledge that NHS trust disclosures often resemble a box-ticking process—predominantly driven by compliance rather than the provision of insightful information. This often results in disclosures that, while meeting regulatory requirements, offer limited interpretive value and pose significant information acquisition costs for external parties. Due to the lack of a continuous transparency variable, we are unable to visually demonstrate these effects, and thus the results should be interpreted with caution.

Table F-1. Predictor construction and the relationship between workforce DEI and patient safety incidents: Non-linear moderation of transparency

Variable	Incidents $_{t+1}$ (1)
Workforce DEI	-112.896*** (20.139)
Transparency	53.286 (102.341)
Transparency^2	-24.206 (46.840)
Workforce DEI × Transparency	181.655*** (38.200)
Workforce DEI × Transparency^2	-64.638*** (16.175)
Admissions	-0.324 (0.294)
SHMI	-0.534 (5.730)
Fundings	0.134 (0.440)
Size of Staff	62.977* (37.145)

Bed Occupancy	14.493 (46.620)
Trust fixed effects	Yes
Year fixed effects	Yes
Adj R^2	0.133
Obs.	467

Appendix G. NHS Responses to Policy Pressure from the UK Government in 2020

	Policy highlights in the Diversity and Inclusion Strategy 2019 to 2023 (Published in 2019)	NHS's responses to the Diversity and Inclusion Strategy (Published in 2020)
Objectives	• Exceed statutory requirements, championing a culture of diversity and inclusion. • Celebrate areas of strong representation while addressing areas that need improvement. • Cultivate an inclusive organizational culture that values diversity in interactions internally and externally. • Integrate diversity and inclusion into our organizational ethos.	• Recruitment across communities to ensure a diverse workforce. • NHS actively working with educational institutions to diversify health and care careers. • Prioritizing behaviors and culture changes. Ensuring a compassionate and inclusive culture.
Statutory compliance	• Committed to upholding the Equality Act, prohibiting discrimination based on age, disability, gender, marital status, maternity, race, religion, sex, and sexual orientation. • The policies will remain updated, inclusive, and compliant. • An emphasis will be placed on transparency, training, and reporting, especially regarding our gender pay gap and public sector equality duty.	• Continuing professional development support, with protective time and supportive supervision. • New funding to support the professional development of nurses, midwives, and allied health professionals. • Expansion of e-learning materials, including simulation.
Representation enhancement	• Transparent reporting on representation, both strengths and areas of improvement. • Active measures to address underrepresentation at all levels. • Commitment to fairness in recruitment and retention, with inclusion as a core tenet. • Focus on expanding representation beyond protected characteristics, including socio-economic factors. • Participation in positive action initiatives such as META, Levelling the Playing Field, and SDIP.	• A blended learning nursing degree program to increase appeal and widen access to a nursing career. • Recruiting and deploying staff across organizations and geographies. • Focus on better use of routes into NHS careers (e.g., volunteering, apprenticeships) • Active work alongside educational institutions and communities to diversify the health and care workforce.
Inclusive culture building	• Inclusivity as a primary objective in our business plans and people strategy. • Raising awareness about diversity benefits, protected characteristics, and potential barriers. • Enhanced visibility of champions and staff networks. • Prioritizing mental health, promoting flexible work arrangements, and ensuring supplier alignment on diversity principles.	• The importance of behavior and culture change, with a strong appetite to do things differently. • The emphasis on a compassionate and inclusive culture. • Focus on improving staff banks' performance and experience. • Development of workforce sharing agreements to enable rapid deployment and flexibility.
Governance and accountability	• All staff share the responsibility for fostering an inclusive environment. • The Directors Group assumes ultimate accountability for the strategy's aims, monitoring its delivery.	• NHS England and NHS Improvement work closely with employers and systems to optimize performance. • Guidelines developed for easy sharing of information, including HR records. • Support for the trial of the COVID-19 digital staff passport.
Implementation and progress monitoring	• While many actions are part of regular operations, specific additional steps will enhance the strategy's effectiveness. • A continually updated implementation action plan will guide the strategy, with annual progress reviews by the Diversity and Inclusion Forum. • Progress measurements will encompass diverse metrics, from HR statistics, gender pay gap data, to benchmarking against broader entities. • Strategy review is scheduled for 2023.	• Establishment of a £10m fund for increased clinical placement capacity. • Guidelines developed by NHS England and NHS Improvement for rapid deployment and staff movement. • Expected further action plans for 2021/22 once funding arrangements are confirmed.