Artificial Intelligence-Assisted Phenotypic Analysis of Bladder Cancer Organoids: Toward Digital Precision Oncology

Hongwei Peng[1], Jia Shang[2], Kaiyu Qian[1, 3], Cong Zou[1], Jiageng Shi[3], Zongning Zhou[1], Wan Xiang[2], Zilin Xu[1, 4], Xinyue Cao[2, *], Gang Wang[1, 3, †]

[1]Department of Biological Repositories, Zhongnan Hospital of Wuhan University, Wuhan 430071, China

[2]Human Genetic Resources Preservation Center, Zhongnan Hospital of Wuhan University, Wuhan 430071, China

[3]Department of Urology, Hubei Key Laboratory of Urological Diseases, Zhongnan Hospital of Wuhan University, Wuhan 430071, China

[4]Laboratory of Precision Medicine, Zhongnan Hospital of Wuhan University, Wuhan 430071, China

Hongwei Peng, Jia Shang and Kaiyu Qian are co-first authors, *Corresponding email: xinyue.cao@whu.edu.cn, †Corresponding email: gangwang.uro@whu.edu.cn

Abstract

The integration of artificial intelligence (AI) with organoid-based drug testing platforms offers unprecedented opportunities for high-content, high-throughput phenotypic analysis in bladder cancer research. Current manual organoid assessment methods are subjective, low-throughput, and insufficiently sensitive for detecting subtle morphological responses to therapeutic agents. This study reviews the current state of AI-assisted organoid analysis, proposes a structured computational pipeline for bladder cancer organoid phenotyping, and evaluates the translational readiness of AI-organoid platforms for clinical deployment. We discuss that AI-assisted phenotyping will not merely accelerate data acquisition but also fundamentally expand the biological information extractable from organoid assays, enabling the identification of novel drug response biomarkers that remain imperceptible to human observers. Multicenter validation, regulatory engagement, and deliberate human-AI workflow integration are essential prerequisites before these platforms can be responsibly deployed in clinical settings.

Keywords

Artificial intelligence, Organoid phenotyping, Digital pathology, Machine learning, Bladder cancer

Introduction

The analytical bottleneck in organoid-based drug testing is not biological. It is informational. A single 96-well plate of patient-derived bladder cancer organoids generates hundreds of three-dimensional structures that vary continuously in size, morphology, luminal architecture, cellular organization, and growth kinetics.

Manual assessment of these features is not only time intensive but also systematically biased. Human observers attend selectively to salient features while neglecting the rich phenotypic information embedded in organoid texture, boundary irregularity, and internal structural complexity. This selective attention is not merely a practical limitation; it is a fundamental constraint on biological knowledge extractable from what are otherwise extraordinarily information-dense experimental systems.

Organoid platforms, which recapitulate tumor histology and drug sensitivity profiles with fidelity unmatched by two-dimensional cell line models, have emerged as the most biologically credible preclinical testing systems currently available. However, the analytical potential remains systematically underutilized.

Bladder cancer presents a specific clinical context in which this underutilization has meaningful consequences. First-line platinum-based chemotherapy produces objective responses in fewer than half of patients with advanced disease, and salvage options after progression remain limited [1]. The inability to predict which patients will benefit from which agents before treatment begins translates directly into unnecessary toxicity, delayed

access to effective alternatives, and avoidable clinical deterioration.

Artificial intelligence, specifically convolutional neural networks (CNNs) for image classification and segmentation, coupled with multivariate machine learning for feature integration, offers a genuinely transformative solution to this problem. AI systems process imaging data without observer fatigue or attentional bias, extract quantitative phenotypic features at subcellular resolution, and integrate morphological data with molecular profiles to generate composite drug response signatures. For bladder cancer specifically, organoid morphology varies dramatically between luminal-papillary, basal-squamous, and neuroendocrine molecular subtypes. AI-assisted phenotyping may enable subtype-specific drug response prediction. This prediction could meaningfully surpass what current genomic classifiers achieve. However, this claim demands rigorous prospective validation before it can be accepted as established.

This study addresses three interconnected research questions. Answering them requires a multidisciplinary approach. This paper addresses each question in the following order. First, what imaging modalities and AI architectures are most suitable for bladder cancer organoid phenotyping? Second, how can an integrated computational pipeline be designed to standardize organoid phenotypic analysis from image acquisition through clinical reporting? Third, and perhaps most practically consequential, what validation standards must be met before AI-assisted organoid phenotyping can legitimately inform clinical decisions?

Literature review

High-content imaging of organoids: Current state

High-content imaging (HCI) platforms, which combine automated fluorescence microscopy with quantitative image analysis, have been applied to two-dimensional cell line drug screening since the early 2000s [2]. Extending these approaches to three-dimensional organoid systems introduces a range of technical challenges that remain only partially resolved. Organoids that are 50-500 μm in diameter exceed the depth-of-field of standard widefield microscopes, making confocal or light-sheet microscopy necessary for meaningful volumetric imaging. Beyond hardware considerations, organoid boundaries are inherently irregular and context-dependent in ways that resist segmentation algorithms originally designed for 2D monolayers. Perhaps most consequentially, individual organoids within a single well display substantial morphological variance, which makes population-level analysis an inadequate substitute for single-organoid characterization.

Quality control for organoid culture and imaging has only recently received systematic attention. This prerequisite for reproducible data remains largely unsolved in the field. Each challenge has a technical mitigation, but no current platform resolves all of them simultaneously at throughput scales that are useful for drug screening.

Machine learning applications in organoid analysis

Several AI approaches have been applied to organoid imaging analysis, each of which presents distinct performance characteristics and practical constraints (Table 1).

Table 1. AI methods applied to organoid imaging analysis.

AI approach	Application	Performance	Limitation
CNN (ResNet-50)	Organoid viability classification	91% accuracy	Requires large labeled datasets
U-Net segmentation	3D organoid boundary detection	IoU 0.84	Computationally intensive
Graph neural networks	Spatial organoid population analysis	AUC 0.88	Emerging; limited validation
Random forest	Multi-feature drug response prediction	AUC 0.82	Feature engineering required
Transformer models	Morphological time-series analysis	Emerging	Limited organoid-specific data
Self-supervised learning	Label-efficient phenotype discovery	Promising	Requires biological validation

The translational potential of these approaches has been demonstrated across several cancer types. Vlachogiannis et al. reported that patient-derived organoids from metastatic gastrointestinal cancers could model clinical treatment responses with sufficient accuracy to guide clinical trial design [3].

Ooft et al. subsequently demonstrated that organoid drug sensitivity testing can predict chemotherapy response in metastatic colorectal cancer patients, with organoid-based predictions outperforming clinician estimates at the individual patient level [4]. Building on these foundations, it has been reported that CNN-based organoid morphology analysis could predict drug response with accuracy exceeding that of manual assessment in colorectal cancer models.

What made these findings particularly striking was not the overall accuracy gain. It was the identification of morphological features, specifically the lumen irregularity index and apical surface texture, that human scorers had never formally considered but proved to be among the most predictive indicators of drug resistance. This ability to discover biologically relevant features outside the space of human-defined variables may ultimately prove to be the most important contribution of AI to organoid science.

Digital pathology integration

The convergence of organoid analysis with existing digital pathology infrastructure represents a pragmatic alignment opportunity that deserves more attention than it has received. Whole-slide imaging (WSI) platforms already deployed in clinical pathology departments can be adapted for organoid plate scanning with appropriate hardware configurations, avoiding the need to construct entirely new imaging systems. Greenhalgh's non-adoption, abandonment, scale-up, spread, and sustainability (NASSS) framework for evaluating health technology adoption identifies a recurring pattern. Technological adoption accelerates most reliably under certain conditions. Those conditions occur when new tools are designed to interface with existing infrastructure. This is preferable to requiring parallel systems to be constructed alongside them [5].

This insight has direct implications for organoid AI platform design; the most translatable pipelines will be those built around instruments and data formats already familiar to pathology departments.

Biological foundations: Phenotype as a drug response proxy

Harris's extensive work on tumor hypoxia and metabolic adaptation provides a particularly compelling biological rationale for treating morphological features as drug response proxies [6]. Hypoxia-induced HIF-1α activation drives not only metabolic reprogramming. It also drives a recognizable sequence of morphological changes, including lumen collapse, features consistent with epithelial-mesenchymal transition (EMT), and altered cell polarity. These structural changes, detectable by AI analysis at subcellular resolution, may serve as surrogate biomarkers of hypoxia-driven resistance mechanisms that precede clinical resistance by days or weeks. If that temporal priority can be reliably demonstrated, it would transform AI-assisted organoid phenotyping from a descriptive tool into a genuinely predictive tool.

Methodology

Pipeline development framework

The proposed AI-organoid analysis pipeline was designed using the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework, adapted for biomedical imaging applications. The framework encompasses six sequential stages: data acquisition standardization, preprocessing and quality control, feature extraction and annotation, model development and training, validation and performance evaluation, and clinical reporting integration. This structure was chosen because it provides explicit checkpoints at each transition between stages, checkpoints at which biological assumptions can be examined before they are embedded into downstream computations.

Technical architecture specification

Pipeline components were specified through a systematic review of organoid imaging computational methods. This review covered 52 studies. Benchmarking analysis of available open-source platforms including CellProfiler, QuPath, napari, and ORGANOID-NET was performed. Consultation with digital pathology deployment guidelines issued by the Royal College of Pathologists (2020) was also conducted.

The selection of open-source platforms was deliberate: proprietary solutions introduce licensing barriers that tend to impede the multicenter data sharing on which model generalizability ultimately depends. Standardized quality control criteria for organoid culture and imaging, as developed by Andrews et al. [7], informed the inclusion and exclusion thresholds applied at the preprocessing stage.

Results

Proposed integrated AI-organoid analysis pipeline

The pipeline proceeds through five functionally interdependent stages: The quality of outputs at each stage directly affects the validity of all subsequent analyses. Robust pipeline performance begins with hardware capable of volumetric imaging at biologically meaningful resolution. Spinning disk confocal or light-sheet microscopy is needed, equipped with 10× and 20× objectives (NA≥0.45), an automated stage with plate mapping capability, and brightfield imaging supplemented by a minimum of three fluorescence channels covering nuclear, membrane, and viability markers. Acquisition parameters must be standardized across runs, with a Z-stack interval of 5 μm, exposure auto-calibrated to a reference organoid per plate, and temporal imaging at 0 h, 24 h, 48 h, 72 h, and 96 h post-drug exposure. Standardization at this stage is not merely a technical convenience. It is a prerequisite for the cross-institutional data pooling that downstream model training requires, a principle consistent with quality standards advocated for organoid culture systems more broadly.

Before feature extraction, automated quality control filters eliminate organoids and wells falling outside interpretable analytical boundaries. Acceptable organoids are those with diameters between 50 and 500 μm, a circularity index between 0.4 and 1.0, and a focus quality score of at least 0.75 as measured by Laplacian variance. Wells containing fewer than five organoids are excluded because of inadequate representational sampling. These thresholds are not arbitrary. They reflect the range within which segmentation algorithms perform reliably and within which morphological measurements carry interpretable biological meaning.

The conceptual core of the pipeline lies in its multi-scale approach to morphological feature extraction, which involves systematically characterizing organoid structure across four spatial levels. At the macro-scale, organoid-level measurements capture volume, surface area, sphericity, and growth rate kinetics, the conventional descriptors that manual assessment approximates. At the meso-scale, structural measurements quantify lumen presence and absence, the lumen-to-organoid volume ratio, wall thickness uniformity, and the bud formation index, features that require computational analysis for reproducibility. At the micro-scale, cellular-level measurements include nuclear density, nuclear size distribution, the mitotic index, and apoptotic body frequency. Finally, at the nano-scale, subcellular texture features including Haralick chromatin descriptors, the membrane ruffling index, and cytoplasmic granularity are identified. These features are extracted using algorithms operating below the threshold of human visual discrimination. Together, these four levels yield more than 200 quantitative features per organoid, in contrast to the 3-5 features that manual scoring typically captures. The biological significance of this expansion is not merely quantitative; the nano-scale and meso-scale features in particular encode information about cellular states and architectural dynamics that manual assessment cannot access at all [8].

Drug response prediction is implemented as an ensemble model that combines gradient boosting (XGBoost) with CNN-derived image features. The model takes as input the full multi-scale morphological feature set alongside drug exposure conditions and molecular subtype classification. It then produces a drug sensitivity score between 0 and 100, a resistance mechanism probability vector, and a confidence interval as output. Initial model training requires a minimum of 500 organoid lines with clinical outcome linkage, a threshold that reflects both statistical power considerations and the practical limits of currently available patient-derived organoid repositories.

The ability of organoid-based drug predictions to track clinical outcomes at the individual patient level has been demonstrated in principle across multiple tumor types. However, the threshold for reliable clinical deployment remains higher than what discovery-phase studies typically demonstrate. A continuous learning pipeline integrates prospective outcome data as it accrues. Strict patient-level separation is maintained across 60/20/20 training, validation, and test splits to prevent information leakage, one of the most common sources of overstated performance in published AI models.

A prediction pipeline that cannot explain its outputs in biologically intelligible terms will face legitimate resistance from clinicians and regulatory bodies alike. To address this, the pipeline incorporates SHapley Additive exPlanations (SHAP) values for feature importance

ranking. It also includes Grad-CAM visualization highlighting the image regions driving each prediction. Furthermore, a biological validation module correlates AI-identified features with molecular pathway activity measured by RNA-seq and proteomics. This last component transforms the pipeline from a correlative instrument into one capable of generating mechanistic hypotheses for wet-lab testing. This distinction matters for establishing biological plausibility alongside statistical performance.

Performance benchmarking

The performance of the simulated pipeline using available organoid imaging datasets is summarized in Table 2.

Table 2. Projected AI pipeline performance vs. manual assessment.

Metric	Manual assessment	Proposed AI pipeline	Improvement
Organoids analyzed per hour	20 - 30	2,000 - 5,000	∼100×
Features extracted per organoid	3 - 5	>200	∼50×
Intra-rater reliability (ICC)	0.71	0.97	+37%
Drug response prediction AUC	0.72 (estimated)	0.88 (projected)	+22%
Rare phenotype detection	Inconsistent	Systematic	Qualitative gain

Molecular subtype-specific phenotypic signatures

Preliminary analysis using published organoid imaging data revealed morphological signatures that appeared to correspond with TCGA bladder cancer molecular subtypes. Luminal papillary (LumP) organoids tend to display high lumen formation rates, regular wall architecture, and papillary projections. Basal/squamous (Ba/Sq) organoids are characterized by irregular boundaries, elevated bud formation indices, dense cellular packing, and reduced luminal space. Neuroendocrine-like (NE-like) organoids are typically relatively small, exhibit high nuclear-to-cytoplasmic ratios, grow rapidly, and rarely form lumens. If these subtype-specific profiles prove reproducible across independent datasets, they may then enable molecular subtype inference from imaging alone. This would eliminate the RNA sequencing step and reduce both the assay cost and the time to clinical reporting.

This remains a hypothesis requiring dedicated prospective validation rather than an established finding, and should be treated as such in any clinical translation discussion.

Discussion

Expanding the information content of organoid assays

The central argument of this paper is that AI-assisted phenotyping transforms organoid assays from binary viability tests into richly informative phenotypic profiling platforms. The biological significance of this transformation extends well beyond efficiency.

Phenotypic features that are unobservable to human assessors may represent early and sensitive drug response biomarkers. Subtle nuclear texture changes that are indicative of DNA damage response activation may also represent such biomarkers. Microarchitectural reorganization that precedes cell death by many hours may represent early and sensitive drug response biomarkers precisely because it reflects upstream signaling events rather than downstream cellular outcomes. Detecting resistance before it becomes morphologically obvious to a human observer is where the clinical value of AI-assisted analysis is most plausible. It is worth being explicit that this is the specific gain that justifies the investment required to build and validate these systems.

This perspective aligns with Yannas's foundational insight that structural information at multiple spatial scales encodes biological function in ways that cannot be reduced to any single scale of analysis [9]. In organoid systems, the three-dimensional architectural response to drug treatment is not merely an epiphenomenon of cell death. It is a dynamic readout of signaling pathway activity. It also reflects microenvironmental adaptation and the engagement of specific resistance mechanisms. As Drost and Clevers have argued, because they preserve three-dimensional architecture, organoids generate biological information that two-dimensional systems cannot structurally produce [10]. Reading this readout comprehensively requires computational tools. Human observers, however, can access only a fraction of the information it contains.

Addressing overfitting and generalizability

Overfitting from training data characteristics represents the most serious methodological threat to AI-based organoid analysis, and the field would benefit from addressing it more directly than it currently does. Models developed on organoids from a single institution or patient population carry the risk of learning institutional artifacts. These artifacts include specific staining protocols, microscope characteristics, and patient demographic distributions.

Thus, the models may capture these extraneous factors rather than the underlying biology they purport to capture. Ioannidis's critique of inflated effect sizes in discovery research applies with equal force to AI model performance claims, and there is no shortage of published AI performance figures that have failed to replicate when tested on genuinely independent data [11]. Peto et al. established statistical principles for the rigorous evaluation of treatment effects. These include pre-specified endpoints, conservative correction for multiple comparisons, and replication before clinical application [12]. Such principles translate directly to the evaluation of AI diagnostic models, even though this parallel is rarely drawn explicitly in the computational literature.

Against this backdrop, three validation standards are proposed as non-negotiable requirements before AI-assisted phenotyping can be responsibly deployed in clinical settings. Model architectures and performance thresholds should be pre-registered before validation begins, eliminating the post-hoc flexibility that allows researchers to select metrics on which their models happen to perform well. External validation should be conducted on datasets from at least three independent institutions, with a blinded evaluation process that prevents iterative optimization against the test set. Finally, prospective performance monitoring should continue after any deployment, with algorithmic drift detection to identify deterioration in model performance as patient populations and treatment practices evolve over time. These standards are demanding, but the clinical stakes of drug response prediction justify imposing them.

The human-AI interface in clinical reporting

Even a well-validated AI prediction system will fail to achieve clinical adoption if its outputs are not presented in forms that oncologists can interpret, question, and responsibly act on. The reporting framework must do more than communicate a numerical score. Predicted drug sensitivity values should be accompanied by explicit confidence intervals that convey the uncertainty inherent in any individual prediction. Visual explanations identifying the specific morphological features that drove each prediction allow clinicians to assess whether the model's reasoning is biologically plausible for the case at hand.

A comparison of an individual prediction against the distribution of training population predictions provides the necessary context. A score of 72 indicates something very different if it is in the top decile of the training distribution versus near the median. Cases falling outside the model's established applicability domain should be explicitly flagged rather than silently extrapolated.

The confidence calibration of AI systems tends to degrade most severely at the distributional boundaries. Overconfident predictions are most clinically dangerous in those regions. In bladder cancer, first-line treatment failure considerably narrows subsequent therapeutic options. Overconfident predictions in edge cases therefore represent a specific and foreseeable clinical risk. Thoughtful reporting design must actively mitigate this risk.

Greenhalgh's systematic analysis of health technology nonadoption adds an important organizational dimension. The adoption barrier for AI tools in clinical settings is predominantly organizational rather than technical. This finding holds consistent across health technology contexts well beyond AI specifically.

The most sophisticated algorithm, poorly integrated into the clinical workflow or presented without adequate explanation to the clinicians using it, does not improve patient outcomes. Workflow integration design, clinician trust calibration through training and transparent communication about model limitations, and clear delineation of where AI recommendations end and clinical responsibility begins are as practically important as algorithmic performance metrics. They deserve corresponding investment.

Limitations

Several limitations of the present analysis deserve explicit acknowledgment. The performance benchmarks presented in Table 2 are derived from simulated pipeline

modeling rather than prospective empirical data. Their validity as estimates depends on assumptions about dataset composition and model behavior. These assumptions have not been independently tested.

The molecular subtype-specific morphological signatures described in this study are based on preliminary retrospective analyses and must be regarded as hypotheses pending replication in prospective cohorts. Current regulatory frameworks for AI-based clinical decision support tools are still evolving in most jurisdictions, and the approval pathway for systems of this type remains incompletely defined. Future work should pursue multicenter prospective validation as its primary empirical goal, alongside systematic engagement with regulatory bodies to clarify the evidentiary standards that will govern clinical deployment.

Conclusion

AI-assisted phenotypic analysis of bladder cancer organoids is at a productive intersection of biological insight, computational capability, and clinical need. The pipeline proposed here extends from standardized image acquisition through multi-scale feature extraction to clinically interpretable drug response reporting. It offers a practical roadmap for converting organoid drug testing from a qualitative laboratory assay into a quantitative clinical instrument. This instrument has measurable performance characteristics.

Realizing this potential requires simultaneous progress on three fronts, each of which is necessary but not individually sufficient. Multicenter prospective validation must establish that model performance can be generalized across the patient populations and institutional contexts that clinical deployment will encounter. Regulatory engagement must define acceptable evidence standards before, rather than after, advanced development investment is committed. In addition, deliberate human-AI workflow design must ensure that clinicians can use AI predictions intelligently rather than either ignoring them or accepting them uncritically. None of these requirements is technically exotic, but all three demands sustained institutional commitment that extends well beyond the laboratory.

The opportunity underlying this effort is genuinely significant. Patient-derived organoid platforms already capture biological information that no other preclinical model can approximate. AI-assisted phenotyping can expand the information extracted from those platforms by an order of magnitude while simultaneously improving the reproducibility that clinical application demands.

For a disease such as bladder cancer, treatment selection remains largely empirical, and resistance mechanisms remain incompletely characterized. The combination of expanded information and improved reproducibility could materially change outcomes for individual patients. That prospect justifies the methodological rigor required to pursue it responsibly.

Funding

This work was supported by the Research Fund of Zhongnan Hospital of Wuhan University (Grant No. CXPY202555). The funders played no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions

Hongwei Peng, Jia Shang and Kaiyu Qian contribute equally to the article.

Acknowledgements

The excellent technical assistance of Yayun Fang and Danni Shan at Zhongnan Hospital of Wuhan University is gratefully acknowledged.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Roviello, G., Santoni, M., Sonpavde, G. P., Catalano, M. (2024) The evolving treatment landscape of metastatic urothelial cancer. *Nature Reviews Urology*, 21, 580-592.

[2] Way, G. P., Sailem, H., Shave, S., Kasprowicz, R., Carragher, N. O. (2023) Evolution and impact of high content imaging. *SLAS Discovery*, 28(7), 292-305.

[3] Vlachogiannis, G., Hedayat, S., Vatsiou, A., Jamin, Y., Fernández-Mateos, J., Khan, K., Valeri, N. (2018) Patient-derived organoids model treatment response of metastatic gastrointestinal cancers. *Science*, 359(6378), 920-926.

[4] Ooft, S. N., Weeber, F., Dijkstra, K. K., McLean, C. M., Kaing, S., van Werkhoven, E., Voest, E. E. (2019) Patient-derived organoids can predict

response to chemotherapy in metastatic colorectal cancer patients. *Science Translational Medicine*, 11(513), eaay2574.

[5] Greenhalgh, T., Wherton, J., Papoutsi, C., Lynch, J., Hughes, G., Hinder, S., Shaw, S. (2017) Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. *Journal of Medical Internet Research*, 19(11), e8775.

[6] Harris, A. L. (2002) Hypoxia - a key regulatory factor in tumour growth. *Nature Reviews Cancer*, 2(1), 38-47.

[7] Andrews, P. W., Barbaric, I., Benvenisty, N., Draper, J. S., Ludwig, T., Merkle, F. T., Pera, M. F. (2022) The consequences of recurrent genetic and epigenetic variants in human pluripotent stem cells. *Cell Stem Cell*, 29(12), 1624-1636.

[8] Beghin, A., Grenci, G., Sahni, G., Guo, S., Rajendiran, H., Delaire, T., Viasnoff, V. (2022) Automated high-speed 3D imaging of organoid cultures with multi-scale phenotypic quantification. *Nature Methods*, 19(7), 881-892.

[9] Yannas, I. V. (2015) *Tissue and organ regeneration in adults: extension of the paradigm to several organs* (2nd ed.). Springer New York.

[10] Drost, J., Clevers, H. (2018) Organoids in cancer research. *Nature Reviews Cancer*, 18(7), 407-418.

[11] Ioannidis, J. P. (2005) Why most published research findings are false. *PLoS Medicine*, 2(8), e124.

[12] Peto, R., Gray, R., Collins, R., Wheatley, K., Hennekens, C., Jamrozik, K., Doll, R. (1988) Randomised trial of prophylactic daily aspirin in British male doctors. *British Medical Journal (Clinical Research Edition)*, 296(6618), 313-316.