clinical trials. One reason that cross-validation can inadvertently result in overfitting the held-out data is that the modeler, through iterative model adjustments, may eventually use all the available data. The issue is likely more widespread than typically acknowledged. For example, a comprehensive review of 116 studies across various psychiatric diagnoses found signs of overfitting specifically in studies with small sample sizes (<50 participants) (*2*). Small sample sizes also cause large variance in cross-validation results, and although these issues are well known in statistics and machine learning, many studies still do not follow best practices to improve the outcomes of cross-validation (*3*).

A reliable way to assure the generalizability of machine learning models lies in validating their predictive accuracy on a truly independent, untouched validation sample, known as out-of-sample validation. Often, this approach is not used in clinical studies owing to the challenges associated with acquiring larger datasets and the need for stringent rules governing data acquisition and usage. However, the study by Chekroud *et al.* adds to a growing body of evidence that underscores the necessity of these more robust validation standards to avoid overly optimistic results from machine learning models that fail to generalize to wider clinical contexts.

Even with models that are properly validated and supported by large sample sizes, attempts to predict the clinical outcome or treatment response for individual patients can be unreliable. In the study by Chekroud *et al.*, even when data from multiple clinical trials were pooled to train the model, its predictions still failed to generalize to a new independent trial. The reasons for this are complex and multifaceted. A primary factor is the inherent heterogeneity in data from clinical populations. This issue is particularly prominent in psychiatric disorders, which are typically defined by sets of symptoms (syndromes). Patients with the same diagnostic label may exhibit vastly different symptom profiles that warrant different treatments. Moreover, identical symptoms in different individuals might have distinct biological underpinnings and thus require different therapeutic strategies (*4*). Basing machine learning models purely on diagnostic labels without taking this type of heterogeneity into account can lead to inaccuracies when predicting effective treatment strategies.

A promising approach to address this challenge is to stratify patients into more precisely defined categories, for example, based on underlying symptom causes. This can be achieved, in part, through the use of theory-driven computational models that aim to describe underlying disease mechanisms, a method gaining traction in the field of computational psychiatry. These models are increasingly being used alongside data-driven machine learning techniques, forming powerful tools to tackle the issue of heterogeneity in patient populations (*5, 6*).

Another form of heterogeneity may stem from systematic differences across studies, locations, or time points. As a result, predictions of machine learning models trained on data from a specific context—a population, country, setting, or time period—might rely on features that are associated but not causally related with a clinical outcome in a given study but are not predictive in other contexts. One way to address this heterogeneity is to pool data across multiple studies and sites.

Unreliable predictions may also be the result of outdated outcome measures. Many existing symptom scores are based on questionnaires that may no longer align with understanding of the disease and potentially lead to inaccurate assessments of treatment response. For example, the positive and negative syndrome scale (PANSS) used in the clinical trials from Chekroud *et al.* is gradually being supplanted by more contemporary assessment tools, specifically in the context of negative symptoms in schizophrenia (*7*). If a questionnaire fails to fully capture the true disease burden, it might not accurately detect genuine improvements resulting from treatment. This discrepancy can lead to misclassification of who has or has not benefited from the treatment, which hinders the accurate training of the machine learning model. Similar to the heterogeneity within diagnostic categories, outcome measures will become more accurate with increasing insight into the underlying disease mechanism.

The challenges of using machine learning to predict individual treatment response in medicine, specifically in the context of psychiatry, stem from a complex interplay of issues related to model validation standards, diagnostic heterogeneity, and the relevance of outcome measures used. Addressing these challenges is essential for impactful clinical research and to enable progression toward effective precision medicine. ■

Robert J. and Nancy D. Carney Institute for Brain Science, Brown University, Providence, RI, USA. Email: frederike_petzschner@brown.edu

REFERENCES AND NOTES

1. A. M. Chekroud *et al.*, *Science* **383**, 164 (2024).
2. T. Wolfers *et al.*, *Neurosci. Biobehav. Rev.* **57**, 328 (2015).
3. G. Varoquaux *et al.*, *Neuroimage* **145** (Pt B), 166 (2017).
4. K. E. Stephan *et al.*, *Lancet Psychiatry* **3**, 77 (2016).
5. M. P. Paulus, W. K. Thompson, *Psychopharmacology* **238**, 1231 (2021).
6. Q. J. M. Huys *et al.*, *Nat. Neurosci.* **19**, 404 (2016).
7. B. Kirkpatrick *et al.*, *Schizophr. Bull.* **32**, 214 (2006).

ACKNOWLEDGMENTS

F.H.P. was supported by the Brainstorm Program at the Robert J. and Nancy D. Carney Institute for Brain Science.

10.1126/science.adm9218

MATERIALS SCIENCE

The bumpy road to friction control

The frictional properties of material interfaces can be rationally designed

By **Viacheslav Slesarenko**[1,2] *and* **Lars Pastewka**[1,2]

Friction controls daily life, often without being noticed. It allows walking without slipping, holds sandcastles together, and determines the perceived cleanliness of hair. Little resistance is desired when pedaling bikes, yet the expectation of pulling the brakes is to stop moving. Overall, machines use 20% of the world's energy production to overcome frictional resistance (*1*). Present-day strategies to tune friction, derived from more than a century of engineering insights, often involve the lubrication of interfaces with oils or greases. On page 200 of this issue, Aymard *et al.* (*2*) report an alternative strategy of rationally designing the frictional properties of interfaces. Their approach to friction control may lead to the development of surfaces that adapt to the environment in real time.

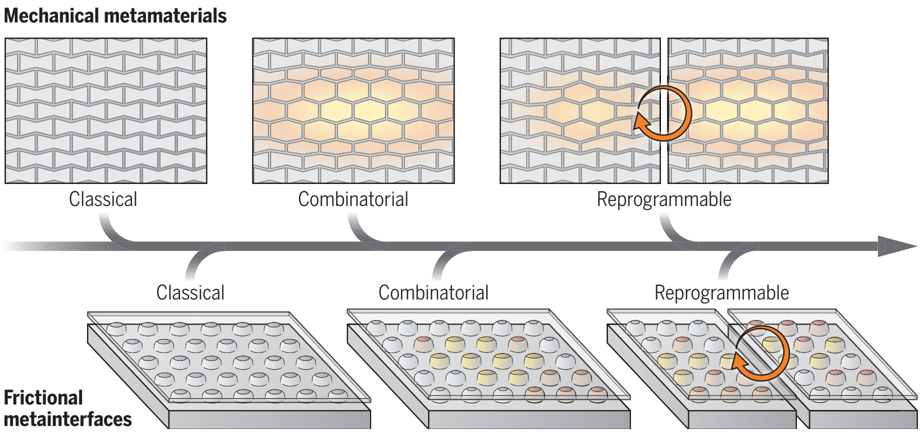
Aymard *et al.* show that small bumps of identical radii (*3*) constitute simple building blocks that can be combined into a frictional metainterface. By using many such bumps on a surface and adjusting their height distribution, the authors could prescribe a desired, even nonlinear, dependence of the frictional force that resists sliding motion on the external load that pushes the sliding interfaces together.

The effect of surface topography on friction has long been known. Charles-Augustin Coulomb, one of the founders of tribology (the science of friction), wrote in 1779 about the interlocking of asperities (*4*), the name given to "bumps" on rough surfaces. Surface topography determines the amount of actual contact that two bodies make. Thus, two bodies typically

[1]Department of Microsystems Engineering, University of Freiburg, 79110 Freiburg, Germany. [2]Cluster of Excellence *liv*MatS, Freiburg Center for Interactive Materials and Bioinspired Technologies, University of Freiburg, 79110 Freiburg, Germany. Email: viacheslav.slesarenko@livmats.uni-freiburg.de; lars.pastewka@imtek.uni-freiburg.de

Metamaterials hint at the future of frictional metainterfaces

Metamaterials progressed from simple repetition of identical unit cells to the assembly of various cells with individually reprogrammable geometries and properties (top). Frictional metainterfaces evolved from simple textures to surfaces with combinatorial topography (bottom). By making each individual bump reprogrammable, adaptive frictional metainterfaces are within reach.



touch only on the highest peaks of their topographies. The contact area of random interfaces increases proportional to the applied load because the surfaces deform to conform (*5*). Friction force is typically proportional to contact area, which is why the friction coefficient—the ratio of friction force to normal force—is constant.

Random roughness is present on all natural and man-made interfaces (*6*), but understanding the influence of roughness on friction is complicated because roughness is scale free; there is no single length scale that stands out. To exert control, engineers have long explicitly introduced geometric structures with well-defined length scales into interfaces. For example, honing produces trenches, and laser patterning produces dimples. Although the effect of these structures on lubricated interfaces can be measured, their functional role is debated, impeding the design of surfaces with desired properties.

"...control of not just simply friction but the whole nonlinear dependency of friction force on normal load appears within reach."

The bumpy metainterfaces of Aymard *et al.* also explicitly introduce a length scale, but in a very controlled manner and for unlubricated contacts. The effects of roughness, mechanics, and chemistry on the frictional properties of the individual bump were experimentally characterized (*3*) and fed into a machine-learned model that was used to design the aggregate response of the metainterface, which consists of many bumps. The height distribution of the bumps was then optimized to yield a desired functional relationship of contact area versus normal load, tuning the frictional response. By this controlled separation of the small-scale and bump-scale response, Aymard *et al.* could "program" the frictional response of a metainterface.

The idea of programming material behavior through geometry and properties of discrete building blocks is ingrained in the field of metamaterials. The first implementation of this concept in mechanics dates back to the 1987 work of the engineering scientist Roderic Lakes, who proposed a material composed of repeating identical unit cells and capable of imitating unusual auxetic behavior that is observed in some foams (*7*). The geometry of the unit cell defines the macroscopic behavior of the entire metamaterial, enabling control over relations between structure and properties (see the figure).

Over time, mechanical metamaterials constructed from a single unit cell gradually morphed into structures that combined unit cells with different geometries and properties within the single metamaterial. Just like Aymard *et al.*'s metasurfaces that have bumps instead of unit cells, this enabled the programming of specific complex mechanical responses through internal organization. The price for versatility is a necessity for advanced modeling methods to search for the structure that facilitates the required behavior. Luckily, this transition nicely coincided with progress in machine learning that has become vital for the efficient inverse design of metamaterials (*8*)—a progress that Aymard *et al.* exploited for designing their metainterfaces. Classical metamaterials can be considered frozen in time and space, unable to change behavior after fabrication. By contrast, building unit cells with geometries that can be altered makes it possible to encode not one but multiple structure-property relationships in the same materials and switch between these relationships on demand (*9*). Such metamaterials might rely on buckling, stimuli-responsive materials or on electromechanical actuators for triggering reprogramming to adjust mechanical behavior (*10*, *11*).

Embedding such active elements into metainterfaces would enable control of friction and facilitate frictional adaptivity. Indeed, simple forms of such reconfigurable interfaces have already been realized to control wettability (*12*). Frictional adaptivity would have numerous applications, such as touch displays with haptic feedback. Current attempts to control friction center around electrochemistry (*13*) or electroadhesion (*14*). A metainterface approach would supersede the chemical route because control of not just simply friction but the whole nonlinear dependency of friction force on normal load appears within reach. Key engineering challenges will revolve around reliability because friction is typically accompanied by wear. Miniaturization of active elements will also be challenging, but either microsystems or stimuli-responsive materials may offer solutions to this. The road toward friction control holds a bumpy, but bright, future. ■

GRAPHIC: N. BURGESS/*SCIENCE*

REFERENCES AND NOTES

1. K. Holmberg, A. Erdemir, *Friction* **5**, 263 (2017).
2. A. Aymard, E. Delplanque, D. Dalmas, J. Scheibert, *Science* **383**, 200 (2024).
3. R. Sahli *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 471 (2018).
4. C. A. Coulomb, *Théorie des Machines Simples* (Imprimerie de Huzard-Courzier, 1821).
5. B. N. J. Persson, *Phys. Rev. Lett.* **87**, 116101 (2001).
6. T. D. B. Jacobs, L. Pastewka, *MRS Bull.* **47**, 1205 (2022).
7. R. Lakes, *Science* **235**, 1038 (1987).
8. S. Bonfanti, R. Guerra, F. Font-Clos, D. Rayneau-Kirkhope, S. Zapperi, *Nat. Commun.* **11**, 4162 (2020).
9. X. Xia, C. M. Spadaccini, J. R. Greer, *Nat. Rev. Mater.* **7**, 683 (2022).
10. H. Yasuda *et al.*, *Nature* **598**, 39 (2021).
11. R. H. Lee, E. A. B. Mulder, J. B. Hopkins, *Sci. Robot.* **7**, eabq7278 (2022).
12. M. Specht, M. Berwind, C. Eberl, *Adv. Eng. Mater.* **23**, 2001037 (2021).
13. F. Bresme, A. A. Kornyshev, S. Perkin, M. Urbakh, *Nat. Mater.* **21**, 848 (2022).
14. M. Ayyildiz, M. Scaraggi, O. Sirin, C. Basdogan, B. N. J. Persson, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 12668 (2018).

ACKNOWLEDGMENTS

The authors receive funding from the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy (grant EXC-2193/1–390951807).

10.1126/science.adn1075