Leveraging agent-based models and deep reinforcement learning to predict taxis in cell migration

Check for updates

Daniel Camacho-Gomez[1] ✉, Raffaele Sentiero[2], Maurizio Ventre[2,3,4] & Jose Manuel Garcia-Aznar[1]

We present a novel computational framework that combines Agent-Based Modeling (ABM) with Reinforcement Learning (RL) using the Double Deep Q-Network (DDQN) algorithm to determine cellular behavior in response to environmental signals. With this approach, the model captures the transduction of environmental cues into biological responses directly from experimental observations, without explicitly predefining cell behavior. This enables the prediction of dynamic, environment-dependent cell behavior and offers a scalable and flexible alternative to traditional rule-based ABM. To illustrate its potential, we present an application to barotactic cell migration data from microfluidic device experiments, where cells adapt their migration behavior based on pressure gradients, demonstrating the model's ability to generalize across varying geometries and pressure configurations. Thus, this approach introduces a novel direction for modeling how cells sense and transduce environmental cues into biological behaviors.

Cell migration is a fundamental biological process involved in development, immune responses, tissue repair, and disease progression[1]. Among the factors that influence this complex behavior, external cues play a crucial role in guiding cells as they navigate their environment. Depending on the nature of these signals, cells exhibit directional migration in response to chemical gradients (chemotaxis), surface-bound cues (haptotaxis), stiffness gradients (durotaxis), and geometrical features (topotaxis)[2]. In addition, pressure gradients represent another type of external signal that can guide cell migration through a process known as barotaxis, which has gained increasing attention in recent years[3]. Barotaxis was first proved in in vitro experiments using microfluidic devices, which showed that cells migrating in confined asymmetric hydraulic environments tend to move toward the path with the least hydraulic resistance[4]. As cells migrate, they displace fluid, generating pressure gradients based on the hydraulic resistance of their surroundings[5]. These pressure gradients, influenced by the asymmetry of the environment, affect decision-making during migration and guide cell movement[6,7].

The influence of pressure gradients on cell migration becomes particularly significant in the context of cancer metastasis, a hallmark of cancer[8] and a leading cause of cancer-related mortality[9]. During metastasis, cells must escape the primary tumor and navigate through the tumor microenvironment (TME)[10], a complex and dynamic space characterized by

profound alterations in biochemical composition, cellular interactions, metabolic state, and mechanical properties[11]. These mechanical alterations include increased matrix stiffness[12,13], compression[14], solid stress[15], and elevated interstitial fluid pressure[16]. As a result, the TME becomes highly confined and dynamic, influencing the presence of pressure gradients that can affect cell migration as they move during metastasis. In this regard[17], studied the migration of cancer cells in asymmetric hydraulic environments, highlighting the preferential migration of cancer cells toward paths with lower hydraulic resistance. Therefore, gaining insight into the mechanisms that drive barotaxis could be crucial for predicting cancer cell metastasis and developing new strategies to disrupt this invasive process.

Computational approaches have become essential in modern biological research due to their ability to model complex biological processes and predict outcomes. Among the various types of models in biology, Agent-Based Models (ABMs) are wide-adopted computational tools that have been widely employed to study cell migration[18–22]. Thus, ABMs allow for the simulation of individual cells and, consequently, investigate the impact of microenvironmental cues on cell behavior[23–25]. Nonetheless, while ABMs provide valuable insights, they need refined calibration techniques to accurately replicate experimental observations.

Machine learning (ML) models are emerging as powerful tools to adapt computational models to data variability, enhancing physics-based models'

[1]Department of Mechanical Engineering, Multiscale in Mechanical and Biological Engineering (M2BE), Aragon Institute of Engineering Research (I3A), University of Zaragoza, Zaragoza, Spain. [2]Department of Chemical, Materials and Industrial Production Engineering. University of Naples Federico II, Naples, Italy. [3]Interdisciplinary Research Centre on Biomaterials. University of Naples Federico II, Naples, Italy. [4]Center for Advanced Biomaterials for Healthcare@CRIB. Fondazione Istituto Italiano di Tecnologia, Naples, Italy. ✉e-mail: dcamacho@unizar.es

ability to capture intricate patterns. When combined with ABMs, ML techniques have the potential to further advance the modeling of cellular behaviors in complex biological systems[26]. Specifically, reinforcement learning (RL) offers powerful frameworks to model complex decision-making processes in dynamic environments[27]. RL enables agents to learn optimal strategies by interacting with the environment, making it highly effective for real-time decision-making tasks. Despite the synergy between ABMs and RL, which has proven to be promising in other fields[28–30] since its introduction by[31], it has yet to be fully applied in biology.

In this work, we aim to predict barotactic cell migration in confined microenvironments through computational modeling. In particular, we present an ABM that integrates an ML model based on RL to predict the influence of pressure gradients on cell migration direction in confined microchannels. To achieve this, we first compute the pressure gradients within the microchannels using Computational Fluid Dynamics (CFD). Then, we simulate cell migration with the ABM, considering the temporal and dynamic variation in the migration direction influenced by the pressure gradients. In this regard, we reproduce the mechanosensing process of cells by adding observation points on the cell membrane where it senses fluid pressure. This sensed pressure is then passed into a Neural Network (NN) that determines the direction of migration. The neural network subsequently adjusts the migration direction based on the pressure gradients sensed by the cell throughout the migration process. To make the cell learn barotactic cell migration decisions, we apply the RL approach, specifically the Double Deep Q-Network (DDQN) algorithm[32]. Finally, we train the model using test geometries and validate it with a real microfluidic geometry designed to study barotaxis from[17], showing agreement between experimental observations and in silico predictions. Thus, this novel approach represents a first step toward building an intelligent in silico cell that reproduces how cells transduce external cues from the environment into migration behaviors.

Results
Computational framework for barotactic cell migration
The proposed computational model has three main parts: the environment, the ABM, and the RL algorithm. The environment's geometry corresponds to the microfluidic device's migration chamber, where the in vitro experiments are conducted[17]. Within this geometry, we perform a CFD simulation to obtain the pressure field $P(\boldsymbol{x})$, which serves as the environmental cue guiding cell migration (Fig. 1A). Then, we simulate cell migration within this environment using an ABM (Fig. 1B). Here, we assume that the temporal variation in the migration direction $e(t)$ is influenced by pressure gradients to replicate barotactic migration. To determine how cell migration is affected by these pressure gradients, we employ a NN trained using an RL approach based on DDQN (Fig. 1C). In our model, the agent observes the fluid pressure at equidistant points around the cell surface, replicating the cell's mechanosensing. This sensed pressure is then passed into the NN, which outputs the probabilities of migration towards discrete action points evenly distributed around the surface. This information is used to update the migration direction at each time step (Δt) throughout the ABM simulation.

Learning barotactic cell migration behavior
First, we train our model to enable the cell to sense pressure stimuli and respond accordingly by regulating its migration direction, thereby learning barotactic cell migration behavior. To achieve this, the cell (agent) is trained in three different geometries, each featuring a bifurcation point with a three-channel intersection. In each geometry, the outlet boundary condition is placed at different locations after the bifurcation point: the straight channel, the top channel, and the bottom channel. These boundary condition locations are designed to create varying pressure gradients, helping the agent learn how to sense and respond to these gradients by migrating in different directions. In each case, the goal position, $\boldsymbol{x}_g$, which influences the reward function, is located at the outlet boundary. We simulate the migrating cell in each of the three geometries and minimize the resulting loss using the DDQN method. Consequently, the reward across all training geometries is represented in each episode, yielding a mean reward of 0.9999 after 7543 episodes (Fig. 2A). As a result, the cell successfully learns to migrate in response to pressure gradients in these three training geometries, moving toward the higher pressure gradient (Fig. 2B). Consequently, as the cell migrates, we observe a decrease in the mean pressure surrounding it (Fig.

Fig. 1 | Overview of the computational model.
A The environment consists of the pressure gradients within the microfluidic device, where in vitro experiments are conducted. **B** The ABM simulates the migrating cell, considering the influence of pressure gradients on the migration direction $e(t)$. **C** The migration direction $e(t)$ is determined through an NN trained using an RL approach based on DDQN. To do that, the NN receives the sensed pressure from the cell and outputs the probabilities of migration toward different directions. Finally, this new migration direction is passed to the ABM, with the process repeated at each time step (Δt) throughout the simulation.
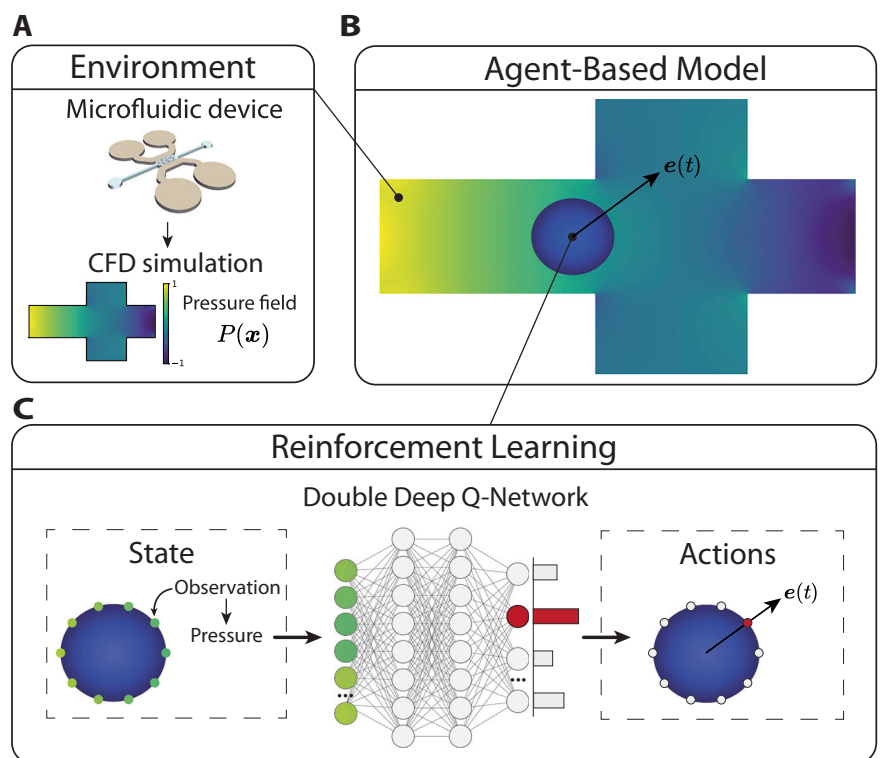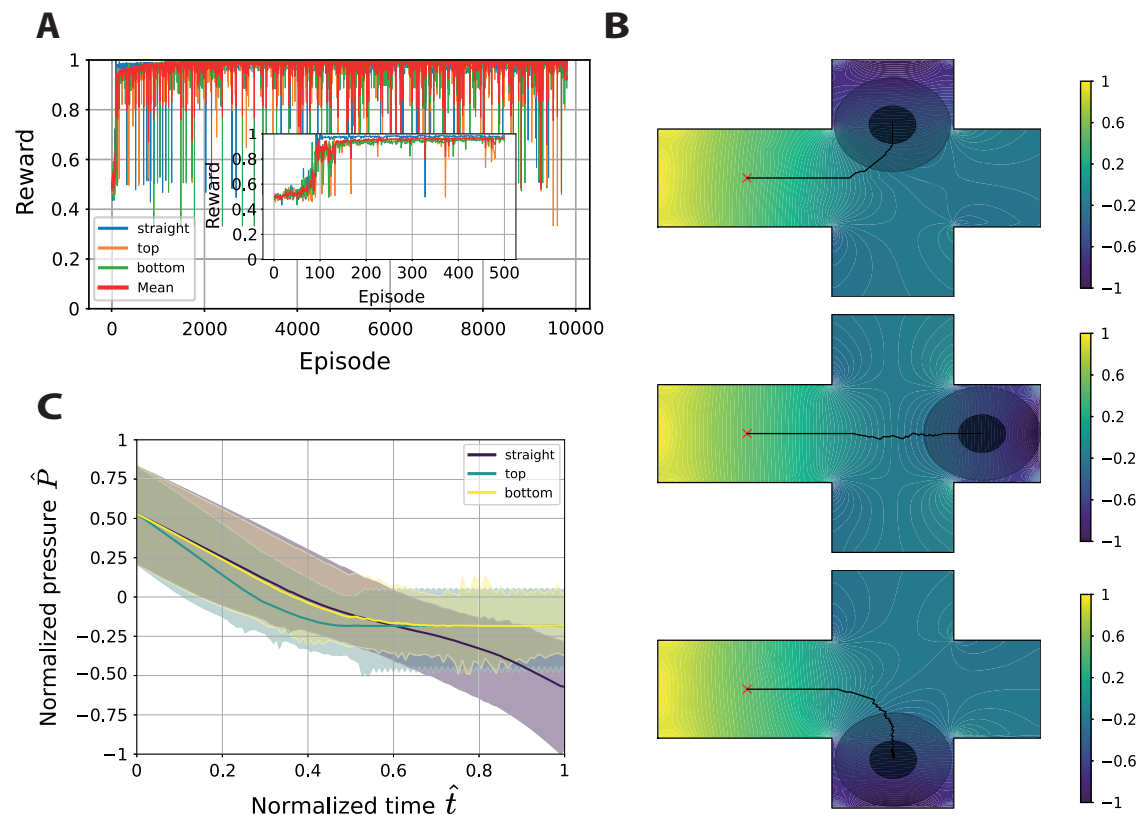
Fig. 2 | Training results for barotactic migration. A Reward across episodes in the straight, top, and bottom geometries, with a zoomed-in detail of the first 500 episodes, reaching a mean value of 0.9999 after 7543 episodes. **B** Trajectories of the cell across the top, straight, and bottom geometries after the training process. **C** Mean normalized pressure around the cell in each geometry, with shaded regions indicating the minimum and maximum values.

Table 1 | Parameters of the computational model

Hyperparameter	Value/Description
Batch size	512 (number of transitions sampled from the replay buffer)
Discount factor (γ)	0.1 (importance of future rewards)
Epsilon (ϵ)	0.9, decays to 0.0001 over 1000 episodes
Learning rate	1×10^{-4} (for the AdamW optimizer)
Soft update rate (τ)	0.005 (rate at which the target network is updated)
Number of observations (N_o)	72 (NN input)
Number of actions (N_a)	16 (NN output)
Time step (Δt)	10 min
Cell velocity (v_c)	0.2 μm/min (mean velocity of the cell from[17])
Radius of the cell (R_c)	10 μm
Radius of the cell nucleus (R_N)	4 μm

2C). The parameters of the DDQN model, along with the parameters of the ABM model, can be found in Table 1.

Predicting in vitro barotactic cell migration

To validate our model, we test its capacity to predict barotactic cell migration in three different real microfluidic devices. In particular, we employ the dead-end, twisted, and tortuous microdevices from previous work[17], replicating the same experimental conditions (Fig. 3). The dead-end microdevice serves as a control device for barotaxis since it maximizes the hydraulic resistance difference between both paths, maximizing the

pressure gradient through the top channel. Consequently, 77% of the cells in the in vitro experiments migrate through top open-path channel, showing a directed migration bias (Fig. 3A, top). Our computational model successfully reproduces barotactic migration, with the cell following the higher pressure gradients to move through the top channel (Fig. 3A, bottom left, and Supplementary Video 1). Thus, we can analyze the pressure sensed by the cell throughout its migration (Fig. 3A, bottom middle). This chart reveals the asymmetrical pressure distribution around the cell and how it progressively decreases over time as the cell moves in the direction of the pressure gradient. In fact, we also compute the mean pressure around the cell throughout its migration, highlighting the cell's tendency to move from regions of higher pressure to lower pressure (Fig. 3A, bottom right).

In the case of the twisted microdevice, in vitro data show that around 65% of cells migrate towards the top path (Fig. 3B, top). Here, the computational model also replicates this cell migration behavior, despite a slight difference in pressure gradients between the bifurcating channels (Fig. 3B, bottom left, and Supplementary Video 2). As a result, the pressure distribution around the cell over time is smoother compared to the dead-end microdevice (Fig. 3B, bottom middle), and the mean pressure profile during migration remains more consistent (Fig. 3B, bottom right).

Regarding the tortuous microdevice, approximately 70% of cells are inclined to migrate toward the bottom tortuous path (Fig. 3C, top). However, there is almost no pressure difference between the bifurcating paths, so our computational model does not predict any barotactic cell decision at the bifurcation point (Fig. 3C, bottom left). Indeed, the pressure distribution around the cell at the bifurcation point is constant, remaining unchanged as the cell does not move from this point (Fig. 3C, bottom middle). Finally, the mean pressure sensed by the cell decreases initially as the cell moves from the starting point to the bifurcation point, and then remains constant, with no differences between mean, maximum, and minimum pressure values,
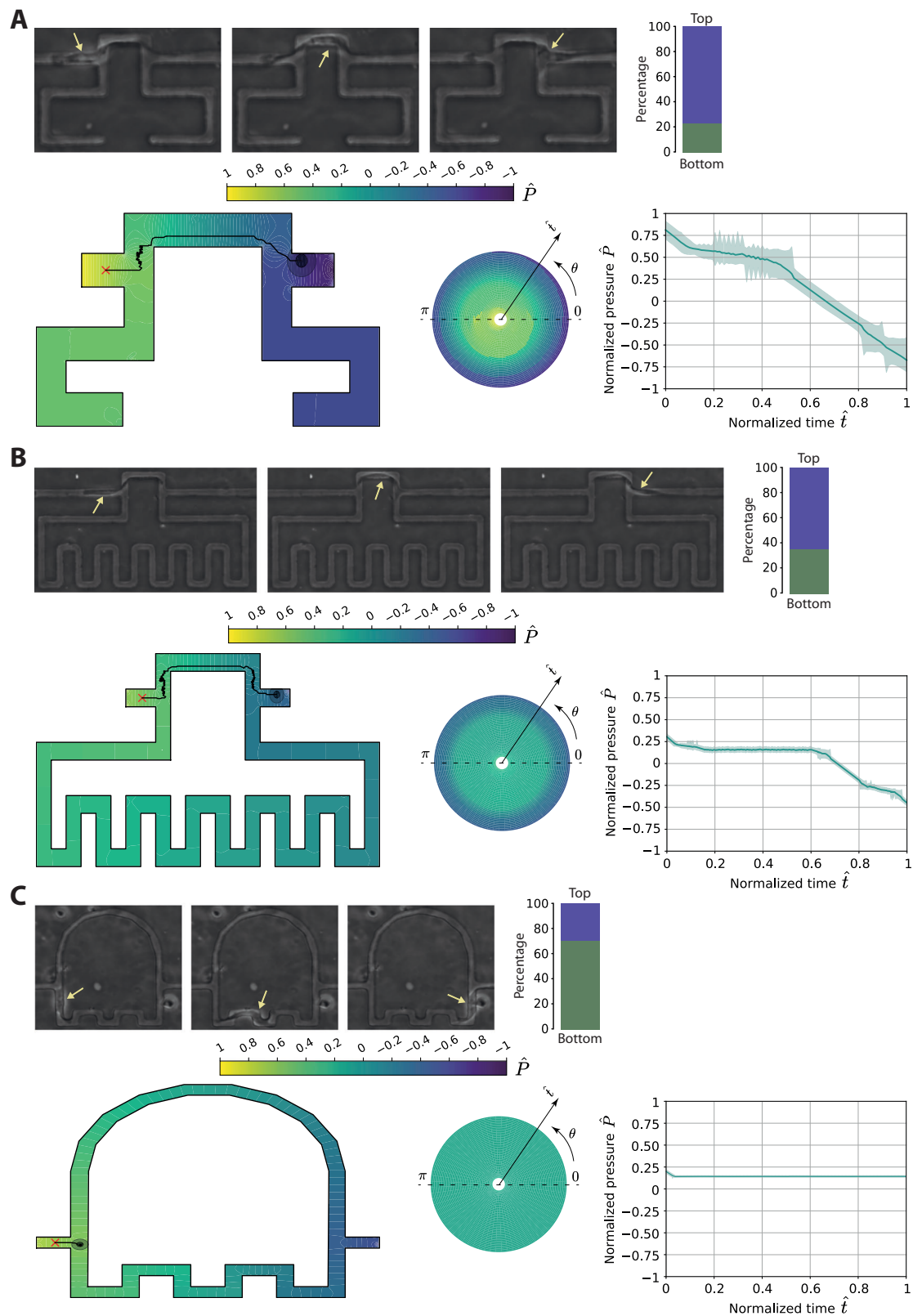
Fig. 3 | Reproducing in vitro barotactic migration. A Dead-end microdevice. B Twisted microdevice. C Tortuous microdevice. For each microdevice, the top panel shows snapshots of the cell migration experiments and the percentage of cells migrating in each channel from[17]. The bottom left panel presents the predicted trajectories. The bottom middle panel displays a donut chart representing the pressure around the cell, where each concentric ring corresponds to the pressure at a given time. The chart includes 50 rings, with the smallest ring representing the initial condition and the largest ring representing the final time. The bottom right panel illustrates the mean normalized pressure around the cell, with shaded regions indicating the minimum and maximum values.

indicating a uniform distribution around the cell. Therefore, these results suggest that barotaxis does not explain the preferential cell migration toward the bottom tortuous path, and other mechanisms might be contributing to this migratory behavior.

Discussion

We presented a novel computational framework to reproduce how cells sense and transduce microenvironmental signals into biological behaviors. To this end, we developed a model that couples ABM with RL using the DDQN algorithm. This integration represents a conceptual advancement in modeling cellular behavior. In contrast to more traditional approaches where cell behavior is pre-programmed or governed by predefined parameters, fixed rules, or simplified decision-making strategies, the RL component enables cells to learn to adapt their behavior based on sensed signals from the microenvironment, directly driven by experimental data. This reduces model bias and allows emergent behavior to arise naturally from the complex interplay of signals influencing cell behavior. It also has the potential to capture the complex, non-linear behaviors characteristic of biological systems by leveraging the power of deep neural networks to model multifaceted biological relationships. Thus, this data-driven approach provides adaptability to dynamic microenvironmental factors, enhancing the model's generalizability under varying conditions and its ability to uncover the underlying mechanisms driving cell behavior. Therefore, it offers a promising framework for better understanding and predicting complex cell behavior in dynamic environments.

We showed an example application to simulate barotactic cell migration, a known phenomenon where pressure gradients guide cell migration in confined environments[4]. To this end, we characterized the microenvironment through the pressure field using CFD, which provides cues that the cell transduces into migration behavior. We then simulated cell migration with a lattice-free ABM, where the direction of migration is regulated by the environmental pressure gradients. To reproduce how cells sense and respond to these varying pressure gradients, we incorporated an ML model. This model consisted of an NN trained through DDQN that receives the pressure distribution around the cell and determines the migration direction.

To validate our model, we applied our framework to real microfluidic devices and tested its performance against experimental in vitro data from the study of [17]. This facilitated the assessment of the model's ability to accurately predict cell migration behavior in experiments, demonstrating its ability to generalize across different geometries and configurations, while reinforcing the robustness and applicability of the approach. Thus, we reproduced barotactic cell migration in the dead-end microdevice, identifying the environmental cues that the cell transduced into biological migration behavior. Interestingly, we observed that the cells oscillate at the bifurcation points, with direction changes that enable them to correctly sense the environmental cues and mechanotransduce this information into migration decisions, as seen in the oscillations of the pressure in Fig. 3A, Supplementary Videos 1 and 2 and in experimental observations[17]. Furthermore, we also predicted barotactic cell migration in the twisted microdevice, demonstrating the model's sensitivity in reproducing barotactic migration despite, slight differences in pressure gradients between bifurcating channels.

However, we did not predict the migration trajectories in the tortuous microdevice. In this microdevice, there are no differences in pressure between the channels at the bifurcation point. Nevertheless, experimental results show a preferential migration toward the bottom tortuous path. Thus, our model indicates that barotaxis is not the driving mechanism guiding cell migration toward this bottom path, suggesting that other factors may be influencing this behavior, as hypothesized in[17]. Possibly, other migration mechanisms may be involved in this directed migration[33], such as the geometric characteristics of the bottom channel, which might provide mechanical support that allows cells to attach to it rather than the curved channel[2]. Therefore, our model serves to predict and identify whether

barotaxis is the underlying mechanism driving cell migration observed in experimental data.

The potential of this approach is not limited to barotactic cell migration. This computational framework can be applied to model directed cell migration in response to other factors, such as chemical gradients (chemotaxis), surface-bound cues (haptotaxis), stiffness gradients (durotaxis), tensile stress (tensotaxis), and geometrical features (topotaxis)[2]. More importantly, in real biological systems, different environmental stimuli often coexist, competing to guide cell migration and potentially outcompeting each other[34]. Thus, this computational framework provides a tool for simulating the combined influence of multiple cues on cell migration. To achieve this, experiments can be conducted in which cells are exposed to these microenvironmental stimuli, and their resulting migration trajectories recorded. The experimental conditions can be simulated in silico to characterize the microenvironment, with a DDQN agent trained and validated using the experimental migration trajectories. Once validated, the model can predict migration patterns under new conditions, including alternative geometries and cue combinations. Additionally, interpretability analyses can quantify the relative contribution of each microenvironmental cue in governing cell behavior, providing mechanistic insights into migration strategies. Therefore, our approach allows learning migration patterns directly from experimental observations, rather than hand-coding them, offering a more comprehensive understanding of directional decision-making in cell migration.

Nonetheless, it is important to acknowledge that our approach relies on certain simplifications, which must be thoroughly justified. The most significant of these relates to the assumption of a time-independent pressure field within the microdevices. In this work, we simulated the fluid flow within the geometry to obtain a static pressure field. However, the actual mechanism involves the cell migrating through the confined channel, pushing the fluid and generating a dynamic pressure gradient as it moves[6]. This effect could be captured by simulating the pressure field resulting from a moving object representing the cell in the CFD simulation. In this case, at each time step, the CFD simulation would be performed with a moving object representing the cell's movement in the ABM. This would enable a fully coupled, real-time simulation between the CFD and the ABM. However, the computational cost of real-time CFD simulations during the training process was prohibitively high. Therefore, we opted to simulate the fluid flow once to characterize the microfluidic device and obtain a representative pressure field.

Another simplification is that we do not explicitly simulate cell deformation. We assumed the cell to be non-deformable since our focus was on reproducing barotactic cell migration trajectories rather than modeling the exact cell shape. To account for cell-wall interactions, we considered the interactions between the wall and the cell nucleus. In this way, we simplified the cell deformation calculation by assuming that the cytoplasm of the cell is deformable and would not oppose resistance, while the nucleus of the cell is non-deformable and resists deformation when it touches the wall. This is consistent with experimental measurements showing that the cell nucleus is stiffer than the cytoplasm[35–37]. Through this approximation, we were able to simulate cell trajectories with a deformable cell, allowing us to focus on the actual trajectories and cell behavior. In this regard, our simulations show that the cell deforms and migrates close to the walls to minimize the traveled distance. This can be observed in Fig. 3A, B (bottom left), where the cell deforms to migrate along the lower wall in the middle section and then moves to the top wall at the outlet. This migration pattern aligns with the in vitro experiments, where cells migrate along the predicted trajectories, attaching to the corresponding walls.

Through this study, we demonstrate that RL can be a powerful tool for reproducing cellular behaviors in ABMs. By leveraging RL, cells can dynamically determine their behavior by transducing external cues in complex microenvironments, reducing reliance on heuristic or rule-based models. Therefore, this approach has the potential to enhance our understanding of how microenvironmental signals influence cell response, contributing to more accurate models of cell behavior.

Methods

Environment

The cell's microenvironment is defined by the pressure field within the microfluidic device. To obtain this pressure field in the microchannels, we perform a Computational Fluid Dynamics (CFD) simulation by means of the Finite Volume Method (FVM) in Ansys Fluent 2023R2. To solve the fluid flow in the microdevice, we use the material properties of liquid water and set an inlet boundary condition with a uniform velocity profile of $0.2\,\mu m/min$, which corresponds to the velocity at which the cell pushes the fluid in experiments[17]. The outlet boundary condition was specified as an atmospheric pressure outlet. We assume laminar flow with a low Reynolds number due to the low velocity magnitude and small characteristic lengths of the microdevices. The solvers employed include a pressure-velocity coupled scheme with a third-order MUSCL scheme for momentum and a second-order scheme for pressure.

Agent-based model

The lattice-free center-based ABM simulates the migrating cell within the microdevice. In this model, we calculate the position of the cell $\boldsymbol{x}_c(t)$ as:

$$\frac{d\boldsymbol{x}_c(t)}{dt} = v_c(\boldsymbol{e}(t;\theta) + \mathcal{H}(R_N - d_w)\boldsymbol{e}_w(t)), \tag{1}$$

with v_c the magnitude of the velocity, $\boldsymbol{e}(t;\theta)$ the unit direction vector of migration, $\mathcal{H}$ is the Heaviside step function, R_N the radius of the cell nucleus, d_w the shortest distance from the cell's position to the wall, and $\boldsymbol{e}_w(t)$ the unit direction vector resulting from the interaction of the cell with the walls. Here, we focus on reproducing pressure-driven cell migration decisions, therefore, we consider a constant velocity magnitude and a variable cell migration direction. The magnitude of the velocity is obtained from the mean velocity of ($v_c = 0.2\,\mu m/min$) measured in the confined experiments in[17]. The direction of migration $\boldsymbol{e}(t;\theta)$ is approximated by a NN with weights θ that takes the local pressure values that the cell senses at each position. Thus, the description of the temporal evolution of $\boldsymbol{e}(t;\theta)$, based on the pressure gradients, replicates barotactic cell migration. The cell-wall interaction is triggered when the cell's nucleus contacts the wall. This interaction is modeled using the Heaviside step function $\mathcal{H}(R_N - d_w)$, which becomes active when the distance d_w becomes less than or equal to R_N. In this case, the unit direction vector $\boldsymbol{e}_w(t)$ opposes the cell's motion, pointing opposite to the vector from the cell center to the collision point, thereby preventing the cell from moving through the wall.

Double deep Q-network

In this study, we employ a Q-learning approach to determine the cell's direction of migration (action a) based on the pressure values sensed by the cell (observation $\hat{o}_t$ at a particular time t) from the pressure field (state s) using OpenAI's Gymnasium library[38]. Q-learning is a reinforcement learning algorithm that aims to estimate the optimal action-value function, $Q(\hat{o}, a)$, which represents the expected cumulative reward of taking an action a based on the observations $\hat{o}$ from the state s, and following an optimal policy thereafter[27]. Specifically, we use Double Deep Q-Network (DDQN) algorithm[32], where a NN approximates the Q-function, $Q(\hat{o}, a; \theta)$, with θ representing the NN's weights. This model mitigates the overestimation bias in action-value predictions from the standard Deep Q-learning by separating the action selection and evaluation processes. DDQN uses two distinct networks: the policy network, which selects actions, and the target network, which evaluates the selected actions, thereby improving stability and learning performance.

The Q-value for a given observation-action pair $(\hat{o}_t, a)$ is approximated as:

$$Q(\hat{o}_t, a; \theta) \approx Q_{\text{policy}}(\hat{o}_t, a), \tag{2}$$

where $\hat{o}_t$ is the observation from the state s made at time t. In our model, the agent observes the fluid pressure values at N_o equidistant points around the surface, replicating the cell's mechanosensing. To ensure consistency, these observed pressure values are normalized to the range $[-1, 1]$ using min-max scaling, where the minimum observed pressure maps to -1 and the maximum to 1. Similarly, the agent's actions correspond to N_a equidistant discrete points around the surface, so the migration direction $\boldsymbol{e}(t;\theta)$ derives from the unit direction vector pointing from the agent's center to the activated action. Thus, N_o and N_a control the spatial resolution of environmental sensing and action selection, respectively, influencing how finely the agent perceives pressure cues and determines its migration direction.

The target Q-value is computed as follows, using the target network and the Double Q-Learning update.

$$\hat{Q}(\hat{o}_t, a_t) = r_t + \gamma \cdot \max_{a'} Q'(\hat{o}_{t+1}, a'; \theta^-). \tag{3}$$

Here, r_t is the reward received after taking action a_t in observation $\hat{o}_t$, γ is the discount factor that weights future rewards, and $Q'(\hat{o}_{t+1}, a'; \theta^-)$ is the Q-value output of the target network for the next observation and a given action a' that maximizes the Q-value. The reward r_t is calculated using the following function:

$$r_t = 1 - \frac{1}{2}\left(\frac{||\boldsymbol{x}_g - \boldsymbol{x}_c(t)||_2}{||\boldsymbol{x}_g||_2}\right)^2, \tag{4}$$

where $\boldsymbol{x}_g$ represents the goal position and $\boldsymbol{x}_c(t)$ is the cell's position. This reward function encourages the cell to move closer to the goal by increasing the reward as the distance decreases, with larger distances penalized more heavily to promote efficient movement.

To balance exploration and exploitation during training, the agent follows an ϵ-greedy policy for action selection. At each time step, it chooses either an action a_t based on the policy network and a random action:

$$a_t = \begin{cases} \arg\max_a Q(\hat{o}_t, a; \theta) & \text{with probability } 1 - \epsilon, \\ \text{random action} & \text{with probability } \epsilon. \end{cases} \tag{5}$$

As training progresses, ϵ decays to encourage greater exploitation of the learned policy.

We use experience replay to break the correlation between consecutive experiences by storing transitions $(\hat{o}_t, a_t, r_t, \hat{o}_{t+1})$ in a replay buffer. During training, random mini-batches of transitions are sampled from the buffer to update the policy network. This improves learning stability by reducing the variance of updates.

The optimization of the policy network is performed by minimizing the Huber loss, a robust loss function for regression tasks, between the Q-values predicted by the policy network and the target Q-values computed using the target network:

$$L(\theta) = \mathbb{E}\left[Huber\left(Q(\hat{o}_t, a_t; \theta) - \hat{Q}(\hat{o}_t, a_t)\right)\right]. \tag{6}$$

Finally, the target network's weights are periodically updated based on the policy network's weights using the soft update rule:

$$\theta^- \leftarrow \tau \cdot \theta + (1 - \tau) \cdot \theta^-, \tag{7}$$

with θ the weights of the policy network, θ^- the weights of the target network, τ the soft update rate.

Data availability

The original codes are publicly available at: https://github.com/daniel-camacho-gomez/CellBarotaxis_DDQN.

Code availability

Original code is publicly available at https://github.com/daniel-camacho-gomez/CellBarotaxis_DDQN.

References

1. Trepat, X., Chen, Z. & Jacobson, K. Cell migration. *Compr. Physiol.* **2**, 2369 (2012).
2. SenGupta, S., Parent, C. A. & Bear, J. E. The principles of directed cell migration. *Nat. Rev. Mol. Cell Biol.* **22**, 529–547 (2021).
3. Lennon-Duménil, A.-M. & Moreau, H. D. Barotaxis: How cells live and move under pressure. *Curr. Opin. Cell Biol.* **72**, 131–136 (2021).
4. Prentice-Mott, H. V. et al. Biased migration of confined neutrophil-like cells in asymmetric hydraulic environments. *Proc. Natl Acad. Sci.* **110**, 21006–21011 (2013).
5. Bruus, H. *Theoretical microfluidics*, vol. 18 https://doi.org/10.1093/oso/9780199235087.001.0001 (Oxford university press, 2007).
6. Stroka, K. M. et al. Water permeation drives tumor cell migration in confined microenvironments. *Cell* **157**, 611–623 (2014).
7. Li, Y., Konstantopoulos, K., Zhao, R., Mori, Y. & Sun, S. X. The importance of water and hydraulic pressure in cell dynamics. *J. cell Sci.* **133**, jcs240341 (2020).
8. Hanahan, D. Hallmarks of cancer: new dimensions. *Cancer Discov.* **12**, 31–46 (2022).
9. Dillekås, H., Rogers, M. S. & Straume, O. Are 90% of deaths from cancer caused by metastases? *Cancer Med.* **8**, 5574–5576 (2019).
10. Balkwill, F. R., Capasso, M. & Hagemann, T. The tumor microenvironment at a glance. *J. cell Sci.* **125**, 5591–5596 (2012).
11. Nagelkerke, A., Bussink, J., Rowan, A. E. & Span, P. N. The mechanical microenvironment in cancer: How physics affects tumours. In *Seminars in cancer biology*, vol. 35, 62–70 https://doi.org/10.1016/j.semcancer.2015.09.001 (Elsevier, 2015).
12. Piersma, B., Hayward, M.-K. & Weaver, V. M. Fibrosis and cancer: A strained relationship. *Biochimica et. Biophysica Acta (BBA)-Rev. Cancer* **1873**, 188356 (2020).
13. Najafi, M., Farhood, B. & Mortezaee, K. Extracellular matrix (ecm) stiffness and degradation as cancer drivers. *J. Cell. Biochem.* **120**, 2782–2790 (2019).
14. Liu, Q., Luo, Q., Ju, Y. & Song, G. Role of the mechanical microenvironment in cancer development and progression. *Cancer Biol. Med.* **17**, 282 (2020).
15. Jain, R. K., Martin, J. D. & Stylianopoulos, T. The role of mechanical forces in tumor growth and therapy. *Annu. Rev. Biomed. Eng.* **16**, 321–346 (2014).
16. Salavati, H., Debbaut, C., Pullens, P. & Ceelen, W. Interstitial fluid pressure as an emerging biomarker in solid tumors. *Biochimica et. Biophysica Acta (BBA)-Rev. Cancer* **1877**, 188792 (2022).
17. Juste-Lanas, Y. et al. Confined cell migration and asymmetric hydraulic environments to evaluate the metastatic potential of cancer cells. *J. Biomech. Eng.* **144**, 074502 (2022).
18. Escribano, J. et al. A hybrid computational model for collective cell durotaxis. *Biomech. modeling Mechanobiol.* **17**, 1037–1052 (2018).
19. Hervas-Raluy, S., Garcia-Aznar, J. & Gomez-Benito, M. Modelling actin polymerization: the effect on confined cell migration. *Biomech. modeling Mechanobiol.* **18**, 1177–1187 (2019).
20. Heck, T. et al. The role of actin protrusion dynamics in cell migration through a degradable viscoelastic extracellular matrix: Insights from a computational model. *PLoS computational Biol.* **16**, e1007250 (2020).
21. Peng, Q., Vermolen, F. J. & Weihs, D. Physical confinement and cell proximity increase cell migration rates and invasiveness: A mathematical model of cancer cell invasion through flexible channels. *J. Mech. Behav. Biomed. Mater.* **142**, 105843 (2023).
22. Camacho-Gomez, D. et al. An agent-based method to estimate 3d cell migration trajectories from 2d measurements: Quantifying and comparing t vs car-t 3d cell migration. *Computer Methods Prog. Biomedicine* **255**, 108331 (2024).
23. Schlüter, D. K., Ramis-Conde, I. & Chaplain, M. A. Computational modeling of single-cell migration: the leading role of extracellular matrix fibers. *Biophysical J.* **103**, 1141–1151 (2012).
24. Daub, J. T. & Merks, R. M. A cell-based model of extracellular-matrix-guided endothelial cell migration during angiogenesis. *Bull. Math. Biol.* **75**, 1377–1399 (2013).
25. Camacho-Gómez, D., García-Aznar, J. M. & Gómez-Benito, M. J. A 3d multi-agent-based model for lumen morphogenesis: the role of the biophysical properties of the extracellular matrix. *Eng. Computers* **38**, 4135–4149 (2022).
26. Camacho-Gomez, D., Sorzabal-Bellido, I., Ortiz-de Solorzano, C., Garcia-Aznar, J. M. & Gomez-Benito, M. J. A hybrid physics-based and data-driven framework for cellular biological systems: Application to the morphogenesis of organoids. *Iscience* **26**, 107164 (2023).
27. Andrew, B. & Richard S, S. *Reinforcement learning: an introduction* (The MIT Press, 2018).
28. Zhu, M., Wang, X. & Wang, Y. Human-like autonomous car-following model with deep reinforcement learning. *Transportation Res. part C: Emerg. Technol.* **97**, 348–368 (2018).
29. Haarnoja, T. et al. Learning to walk via deep reinforcement learning. arXiv preprint arXiv:1812.11103 https://doi.org/10.48550/arXiv.1812.11103 (2018).
30. Zhang, W., Valencia, A. & Chang, N.-B. Synergistic integration between machine learning and agent-based modeling: A multidisciplinary review. *IEEE Trans. Neural Netw. Learn. Syst.* **34**, 2170–2190 (2021).
31. Mnih, V. et al. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602* https://doi.org/10.48550/arXiv.1312.5602 (2013).
32. Van Hasselt, H., Guez, A. & Silver, D. Deep reinforcement learning with double q-learning. In *Proceedings of the AAAI conference on artificial intelligence*, vol. 30 (2016). https://doi.org/10.1609/aaai.v30i1.10295.
33. Paul, C. D., Mistriotis, P. & Konstantopoulos, K. Cancer cell motility: lessons from migration in confined spaces. *Nat. Rev. cancer* **17**, 131–140 (2017).
34. Belotti, Y., McGloin, D. & Weijer, C. J. Analysis of barotactic and chemotactic guidance cues on directional decision-making of dictyostelium discoideum cells in confined environments. *Proc. Natl Acad. Sci.* **117**, 25553–25559 (2020).
35. Caille, N., Thoumine, O., Tardy, Y. & Meister, J.-J. Contribution of the nucleus to the mechanical properties of endothelial cells. *J. Biomech.* **35**, 177–187 (2002).
36. Guilak, F., Tedrow, J. R. & Burgkart, R. Viscoelastic properties of the cell nucleus. *Biochemical biophysical Res. Commun.* **269**, 781–786 (2000).
37. Janota, C. S., Calero-Cuenca, F. J. & Gomes, E. R. The role of the cell nucleus in mechanotransduction. *Curr. Opin. cell Biol.* **63**, 204–211 (2020).
38. Towers, M. et al. Gymnasium: A standard interface for reinforcement learning environments. *arXiv preprint arXiv:2407.17032* https://doi.org/10.48550/arXiv.2407.17032 (2024).

Author contributions

D.C.G. designed the study, implemented the framework, performed all coding, conducted the simulations and analysis, prepared the figures, and wrote the original manuscript. R.A. performed all CFD simulations, prepared

the data, and reviewed the manuscript. M.V. supervised the work and reviewed the manuscript. J.M.G.A. supervised the work, reviewed the manuscript, and secured the necessary resources. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41540-025-00576-0.

Correspondence and requests for materials should be addressed to Daniel Camacho-Gomez.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.