

# Online Research @ Cardiff

This is an Open Access document downloaded from ORCA, Cardiff University's institutional repository: <https://orca.cardiff.ac.uk/id/eprint/158389/>

This is the author's version of a work that was submitted to / accepted for publication.

Citation for final published version:

Khadhour, Sinan, Gallagher, Kevin M., MacKenzie, Kenneth R., Shah, Taimur T., Gao, Chuanyu, Moore, Sacha, Zimmermann, Eleanor F., Edison, Eric, Jefferies, Matthew, Nambiar, Arjun, Anbarasan, Thineskrishna, Mannas, Miles P., Lee, Taeweon, Marra, Giancarlo, Gómez Rivas, Juan, Marcq, Gautier, Assmus, Mark A., Uçar, Taha, Claps, Francesco, Boltri, Matteo, La Montagna, Giuseppe, Burnhope, Tara, Nkwam, Nkwam, Austin, Tomas, Boxall, Nicholas E., Downey, Alison P., Sukhu, Troy A., Antón-Juanilla, Marta, Rai, Sonpreet, Chin, Yew-Fung, Moore, Madeline, Drake, Tamsin, Green, James S.A., Goulao, Beatriz, MacLennan, Graeme, Nielsen, Matthew, McGrath, John S., Kasivisvanathan, Veeru, Chaudry, Aasem, Sharma, Abhishek, Bennett, Adam, Ahmad, Adnan, Abroaf, Ahmed, Suliman, Ahmed Musa, Lloyd, Aimee, McKay, Alastair, Wong, Albert, Silva, Alberto, Schneider, Alexandre, MacKay, Alison, Knight, Allen, Grigorakis, Alkiviadis, Bdesha, Amar, Nagle, Amy, Cebola, Ana, Dhanasekaran, Ananda Kumar, Kond?a, Andra?, Barcelos, André, Galosi, Andrea Benedetto, Ebur, Andrea, Minervini, Andrea, Russell, Andrew, Webb, Andrew, de Jalón, Ángel García, Desai, Ankit, Czech, Anna Katarzyna, Mainwaring, Anna, Adimonye, Anthony, Das, Arighno, Figueiredo, Arnaldo, Villers, Arnould, Leminski, Artur, Chippagiri, Arvinda, Lal, Asim Ahmed, Y?ld?r?m, As?f, Voulgaris, Athanasios Marios, Uzan, Audrey, Oo, Aye Moh Moh, Younis, Ayman, Zelhof, Bachar, Mukhtar, Bashir, Ayres, Ben, Challacombe, Ben, Sherwood, Benedict, Ristau, Benjamin, Lai, Billy, Nellensteijn, Brechtje, Schreiter, Brielle, Trombetta, Carlo, Dowling, Catherine, Hobbs, Catherine, Benitez, Cayo Augusto Estigarribia, Lebacle, Cédric, Ho, Cherrie Wing Yin, Ng, Chi-Fai, Mount, Chloe, Lam, Chon Meng, Blick, Chris, Brown, Christian, Gallegos, Christopher, Higgs, Claire, Browne, Clíodhna, McCann, Conor, Plaza Alonso, Cristina, Beder, Daniel, Cohen, Daniel, Gordon, Daniel, Wilby, Daniel, Gordon, Danny, Hrouda, David, Lau, David Hua Wu, Karsza, Dávid, Mak, David, Martin-Way, David, Suthaharan, Denula, Patel, Dhruv, Carrion, Diego M, Nyanhongo, Donald, Bass, Edward, Mains, Edward, Chau, Edwin, Canelon Castillo, Elba, Day, Elizabeth, Desouky, Elsayed, Gaines, Emily, Papworth, Emma, Yuruk, Emrah, Kilic, Enes, Dinneen, Eoin, Palagonia, Erika, Xylinas, Evangelos, Khawaja, Faizan, Cimarra, Fernando, Bardet, Florian, Kum, Francesca, Peters, Francesca, Kovács, Gábor, Tanasescu, Geroge, Hellawell, Tasso, Giovanni, Lam, Gitte, La Montagna, Giuseppe, Pizzuto, Giuseppe, Gordan, MacLennan, Graeme, Özgür, Günal, Bi, Hai, Lyons, Hannah, arre, Hannah, Ahmed, Hashim, Simpson, Helen, Burden, Helena, Gresty,



Helena, Rios Pita, Hernado, Clarke, Holly, Serag, Hosam, Kynaston, Howard, Crawford-Smith, Hugh, Mostafid, Hugh, Otaola-Arca, Hugo, Koo, Hui Fen, Ibrahim, Ibrahim, Ouzaid, Idir, Puche-Sanz, Ignacio, Toma?kovi?, Igor, Tinay, Ilker, Sahibzada, Iqbal, Thangasamy, Isaac, Cadena, Iván Revelo, Irani, Jacques, Udzik, Jakub, Brittain, James, Catto, James, Green, James, Tweedle, James, Hernando, Jamie Borrego, Leask, Jamie, Kalsi, Jas, Frankel, Jason, Toniolo, Jason, Raman, Jay D., Courcier, Jean, Kumaradeevan, Jeevan, Clark, Jennifer, Jones, Jennifer, Teoh, Jeremy Yuen-Chun, Iacovou, John, Kelly, John, Selph, John P., Aning, Jonathan, Deeks, Jon, Cobley, Jonathan, Olivier, Jonathan, Maw, Jonny, Herranz-Yagüe, José Antonio, Nolzco, Jose Ignacio, Cózar-Olmo, Jose Manuel, Bagley, Joseph, Jelski, Joseph, Norris, Joseph, Testa, Joseph, Meeks, Joshua, Hernandez, Juan, Vásquez, Juan Luis, Randhawa, Karen, Dhera, Karishma, Gronostaj, Katarzyna, Houlton, Kathleen, Lehman, Kathleen, Gillams, Kathryn, Adasonla, Kelvin, Brown, Kevin, Murtagh, Kevin, Mistry, Kiki, Davenport, Kim, Kitamura, Kosuke, Derbyshire, Laura, Clarke, Laurence, Morton, Lawrie, Martinez, Levin, Goldsmith, Louise, Paramore, Louise, Cormier, Luc, Dell'Atti, Lucio, Simmons, Lucy, Martinez-Piñeiro, Luis, Rico, Luis, Chan, Luke, Forster, Luke, Ma, Lulin, Moore, Madeline, Gallego, Maria Camacho, Freire, Maria José, Emberton, Mark, Feneley, Mark, Antón-Juanilla, Marta, Rivero, Marta Viridiana Muñoz, Pir?a, Matea, Tallè, Matteo, Crockett, Matthew, Liew, Matthew, Trail, Matthew, Peters, Max, Cooper, Meghan, Kulkarni, Meghana, Ager, Michael, He, Ming, Li, Mo, Omran Breish, Mohamed, Tarin, Mohamed, Aldiwani, Mohammed, Matanhelia, Mudit, Pasha, Muhammad, Akal?n, Mustafa Kaan, Abdullah, Nasreen, Hale, Nathan, Gadiyar, Neha, Kocher, Neil, Bullock, Nicholas, Campaign, Nicholas, Pavan, Nicola, Al-Ibraheem, Nihad, Bhatt, Nikita, Bedi, Nishant, Shrotri, Nitin, Lobo, Niyati, Balderas, Olga, Kouli, Omar, Capoun, Otakar, Oteo Manjavacas, Pablo, Gontero, Paolo, Mariappan, Paramanathan, Marchiñena, Patricio Garcia, Erotocritou, Paul, Sweeney, Paul, Planelles, Paula, Acher, Peter, Black, Peter C., Osei-Bonsu, Peter K, Østergren, Peter, Smith, Peter, Willemse, Peter-Paul Michiel, Chlosta, Piotr L., Ul Ain, Qurrat, Barratt, Rachel, Esler, Rachel, Khalid, Raihan, Hsu, Ray, Stamirowski, Remigiusz, Mangat, Reshma, Cruz, Ricardo, Ellis, Ricky, Adams, Robert, Hessell, Robert, Oomen, Robert J.A., McConkey, Robert, Ritchie, Robert, Jarimba, Roberto, Chahal, Rohit, Andres, Rosado Mario, Hawkins, Rosalyn, David, Rotimi, Manecksha, Rustom P., Agrawal, Sachin, Hamid, Syed Sami, Deem, Samuel, Goonewardene, Sanchia, Swami, Satchi Kuchibhotla, Hori, Satoshi, Khan, Shahid, Mohammud Inder, Shakeel, Sangaralingam, Shanthi, Marathe, Shekhar, Raveenthiran, Sheliyan, Horie, Shigeo, Sengupta, Shomik, Parson, Sian, Parker, Sidney, Hawlina, Simon, Williams, Simon, Mazzoli, Simone, Grzegorz Kata, Slawomir, Pinheiro Lopes, Sofia, Ramos, Sónia, Rai, Sonpreet, Rintoul-Hoad, Sophie, O'Meara, Sorchia, Morris, Steve, Turner, Stacey, Venturini, Stefano, Almpanis, Stephanos, Joniau, Steven, Jain, Sunjay, Mallett, Susan, Nikles, Sven, Shahzad, Yan, Sylvia, Lee, Taeweon, Uçar, Taha, Drake, Tamsin, Toma, Tarq, Cabañuz Plo, Teresa, Bonnin, Thierry, Muilwijk, Tim, Wollin, Tim, Chu, Timothy Shun Man, Appanna, Timson, Brophy, Tom, Ellul, Tom, Austin, Tomas, Smrkolj, Toma?, Rowe, Tracey, Sukhu, Troy, Patel, Trushar, Garg, Tullika, Ça?kurlu, Turhan, Bele, Uros, Haroon, Usman, Crespo-Atín, Víctor, Parejo Cortes, Victor, Capapé Poves, Victoria, Gnanapragasam, Vincent, Gauhar, Vineet, During, Vinnie, Kumar, Vivek, Fiala, Vojtech, Mahmalji, Wasim, Lam, Wayne, Fung Chin, Yew, Filtekin, Yigit, Chyn Phan, Yih, Ibrahim, Youssed, Glaser, Zachary A, Abiddin, Zainal Adwin, Qin, Zijian, Zotter, Zsuzsanna and Zainuddin, Zulkifli

2022. Developing a Diagnostic Multivariable Prediction Model for Urinary Tract Cancer in Patients Referred with Haematuria: Results from the IDENTIFY Collaborative Study. European Urology Focus 8 (6) , pp. 1673-1682. 10.1016/j.euf.2022.06.001 file

Publishers page: <http://dx.doi.org/10.1016/j.euf.2022.06.001>  
<<http://dx.doi.org/10.1016/j.euf.2022.06.001>>

Please note:

Changes made as a result of publishing processes such as copy-editing, formatting and page numbers may not be reflected in this version. For the definitive version of this publication, please refer to the published source. You are advised to consult the publisher's version if you wish to cite this paper.

This version is being made available in accordance with publisher policies.

See

<http://orca.cf.ac.uk/policies.html> for usage policies. Copyright and moral rights for publications made available in ORCA are retained by the copyright holders.



## Urothelial Cancer

## Developing a Diagnostic Multivariable Prediction Model for Urinary Tract Cancer in Patients Referred with Haematuria: Results from the IDENTIFY Collaborative Study

Sinan Khadhour<sup>a,b,c,\*</sup>, Kevin M. Gallagher<sup>c,d,e</sup>, Kenneth R. MacKenzie<sup>c,f</sup>, Taimur T. Shah<sup>c,g,h</sup>, Chuanyu Gao<sup>c,i</sup>, Sacha Moore<sup>c,j</sup>, Eleanor F. Zimmermann<sup>c,k</sup>, Eric Edison<sup>c,l</sup>, Matthew Jefferies<sup>c,m,n</sup>, Arjun Nambiar<sup>c,f</sup>, Thineskrishna Anbarasan<sup>c,d</sup>, Miles P. Mannas<sup>o</sup>, Taeweon Lee<sup>o</sup>, Giancarlo Marra<sup>p,q</sup>, Juan Gómez Rivas<sup>r</sup>, Gautier Marcq<sup>s,t</sup>, Mark A. Assmus<sup>u</sup>, Taha Uçar<sup>v</sup>, Francesco Claps<sup>w</sup>, Matteo Boltri<sup>w</sup>, Giuseppe La Montagna<sup>w</sup>, Tara Burnhope<sup>x</sup>, Nkwam Nkwam<sup>x</sup>, Tomas Austin<sup>y</sup>, Nicholas E. Boxall<sup>z</sup>, Alison P. Downey<sup>aa</sup>, Troy A. Sukhu<sup>bb</sup>, Marta Antón-Juanilla<sup>cc</sup>, Sonpreet Rai<sup>dd</sup>, Yew-Fung Chin<sup>b</sup>, Madeline Moore<sup>x</sup>, Tamsin Drake<sup>ee</sup>, James S.A. Green<sup>l,ff</sup>, Beatriz Goulao<sup>a</sup>, Graeme MacLennan<sup>gg</sup>, Matthew Nielsen<sup>bb</sup>, John S. McGrath<sup>hh,ii</sup>, Veeru Kasivisvanathan<sup>c,jj,kk</sup>, IDENTIFY Study Group<sup>1</sup>

<sup>a</sup> Health Services Research Unit, University of Aberdeen, Aberdeen, UK; <sup>b</sup> Aberdeen Royal Infirmary, Aberdeen, UK; <sup>c</sup> British Urology Researchers in Surgical Training (BURST) Collaborative, UK; <sup>d</sup> Western General Hospital, Edinburgh, UK; <sup>e</sup> Department of Clinical Surgery, University of Edinburgh, Edinburgh, UK; <sup>f</sup> Freeman Hospital, Newcastle Upon Tyne, UK; <sup>g</sup> Department of Surgery and Cancer, Charing Cross Hospital, Imperial College Healthcare NHS Trust, London, UK; <sup>h</sup> Division of Surgery, Department of Surgery and Cancer, Imperial College London, London, UK; <sup>i</sup> Addenbrookes Hospital, Cambridge, UK; <sup>j</sup> Wrexham Maelor Hospital, Wrexham, UK; <sup>k</sup> Torbay and South Devon NHS Foundation Trust, Torbay, UK; <sup>l</sup> Department of Urology, Whipps Cross Hospital, Barts Health NHS Trust, London, UK; <sup>m</sup> Morrision Hospital, Swansea, UK; <sup>n</sup> Swansea University, Swansea, UK; <sup>o</sup> Department of Urologic Sciences, University of British Columbia, Vancouver, Canada; <sup>p</sup> Department of Surgical Sciences, Città della Salute e della Scienza, Turin, Italy; <sup>q</sup> University of Turin, Turin, Italy; <sup>r</sup> Department of Urology, La Paz University Hospital, Madrid, Spain; <sup>s</sup> Urology Department, Claude Huriez Hospital, CHU Lille, Lille, France; <sup>t</sup> CNRS, Inserm, CHU Lille, Institut Pasteur de Lille, UMR9020-U1277 – CANTHER – Cancer Heterogeneity Plasticity and Resistance to Therapies, University Lille, Lille, France; <sup>u</sup> Division of Urology, Department of Surgery, University of Alberta, Edmonton, Alberta, Canada; <sup>v</sup> Department of Urology, Istanbul Medeniyet University, Istanbul, Turkey; <sup>w</sup> Urological Clinic, Department of Medicine, Surgery and Health Sciences, University of Trieste, Trieste, Italy; <sup>x</sup> University Hospitals of Derby and Burton NHS Foundation Trust, Derby, UK; <sup>y</sup> Department of Urology, Queen Alexandra Hospital, Portsmouth, UK; <sup>z</sup> Salford Royal NHS Foundation Trust, Salford, UK; <sup>aa</sup> Doncaster Royal Infirmary, Doncaster, UK; <sup>bb</sup> University of North Carolina Hospitals, Chapel Hill, NC, USA; <sup>cc</sup> Department of Urology, Hospital Universitario Cruces, Barakaldo, Spain; <sup>dd</sup> St James University Hospital, Leeds Teaching Hospital NHS Trust, Leeds, UK; <sup>ee</sup> The Royal Bournemouth Hospital, Bournemouth, UK; <sup>ff</sup> Healthcare and Population Research, Kings College, London, UK; <sup>gg</sup> Centre for Healthcare Randomised Trials, University of Aberdeen, Aberdeen, UK; <sup>hh</sup> University of Exeter Medical School, Exeter, UK; <sup>ii</sup> Royal Devon and Exeter NHS Foundation Trust, Exeter, UK; <sup>jj</sup> Division of Surgery and Interventional Science, University College London, London, UK; <sup>kk</sup> Department of Urology, University College London Hospitals NHS Foundation Trust, London, UK

<sup>1</sup> IDENTIFY Study Group—PubMed indexed collaborators: Aasem Chaudry, Abhishek Sharma, Adam Bennett, Adnan Ahmad, Ahmed Abroaf, Ahmed Musa Suliman, Aimee Lloyd, Alastair McKay, Albert Wong, Alberto Silva, Alexandre Schneider, Alison MacKay, Allen Knight, Alkiviadis Grigorakis, Amar Bdesha, Amy Nagle, Ana Cebola, Ananda Kumar Dhanasekaran, Andraž Kondža, André Barcelos, Andrea Benedetto Galosi, Andrea Ebur, Andrea Minervini, Andrew Russell, Andrew Webb, Ángel García de Jalón, Ankit Desai, Anna Katarzyna Czech, Anna Mainwaring, Anthony Adimonye, Arighno Das, Arnaldo Figueiredo, Arnauld Villers, Artur Leminski, Arvinda Chippagiri, Asim Ahmed Lal, Asif Yıldırım, Athanasios Marios Voulgaris, Audrey Uzan, Aye Moh Moh Oo, Ayman Younis, Bachar Zelhof, Bashir Mukhtar, Ben Ayres, Ben Challacombe, Benedict Sherwood, Benjamin Ristau, Billy Lai, Brechtje Nellensteijn, Brielle Schreier, Carlo Trombetta, Catherine Dowling, Catherine Hobbs, Cayo Augusto Estigarribia Benitez, Cédric Lebacle, Cherrie Wing Yin Ho, Chi-Fai Ng, Chloe Mount, \* Corresponding author. Health Services Research Unit, University of Aberdeen, Aberdeen, UK. E-mail address: [sinan.khadhour@doctors.org.uk](mailto:sinan.khadhour@doctors.org.uk) (S. Khadhour).

## Article info

### Article history:

Accepted June 4, 2022

Available online 25 June 2022

Associate Editor: Christian Gratzke

### Keywords:

Haematuria

Urinary tract cancer

Urothelial cancer

Bladder cancer

Renal cancer

Prostate cancer

Risk factors

Risk Calculator

## Abstract

**Background:** Patient factors associated with urinary tract cancer can be used to risk stratify patients referred with haematuria, prioritising those with a higher risk of cancer for prompt investigation.

**Objective:** To develop a prediction model for urinary tract cancer in patients referred with haematuria.

**Design, setting, and participants:** A prospective observational study was conducted in 10 282 patients from 110 hospitals across 26 countries, aged  $\geq 16$  yr and referred to secondary care with haematuria. Patients with a known or previous urological malignancy were excluded.

**Outcome measurements and statistical analysis:** The primary outcomes were the presence or absence of urinary tract cancer (bladder cancer, upper tract urothelial cancer [UTUC], and renal cancer). Mixed-effect multivariable logistic regression was performed with site and country as random effects and clinically important patient-level candidate predictors, chosen *a priori*, as fixed effects. Predictors were selected primarily using clinical reasoning, in addition to backward stepwise selection. Calibration and discrimination were calculated, and bootstrap validation was performed to calculate optimism.

**Results and limitations:** The unadjusted prevalence was 17.2% ( $n = 1763$ ) for bladder cancer, 1.20% ( $n = 123$ ) for UTUC, and 1.00% ( $n = 103$ ) for renal cancer. The final model included predictors of increased risk (visible haematuria, age, smoking history, male sex, and family history) and reduced risk (previous haematuria investigations, urinary tract infection, dysuria/suprapubic pain, anticoagulation, catheter use, and previous pelvic radiotherapy). The area under the receiver operating characteristic curve of the final model was 0.86 (95% confidence interval 0.85–0.87). The model is limited to patients without previous urological malignancy.

**Conclusions:** This cancer prediction model is the first to consider established and novel urinary tract cancer diagnostic markers. It can be used in secondary care for risk stratifying patients and aid the clinician's decision-making process in prioritising patients for investigation.

**Patient summary:** We have developed a tool that uses a person's characteristics to determine the risk of cancer if that person develops blood in the urine (haematuria). This can be used to help prioritise patients for further investigation.

© 2022 The Authors. Published by Elsevier B.V. on behalf of European Association of Urology. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Chon Meng Lam, Chris Blick, Christian Brown, Christopher Gallegos, Claire Higgs, Clíodhna Browne, Conor McCann, Cristina Plaza Alonso, Daniel Beder, Daniel Cohen, Daniel Gordon, Daniel Wilby, Danny Gordon, David Hrouda, David Hua Wu Lau, Dávid Karsza, David Mak, David Martin-Way, Denula Suthaharan, Dhruv Patel, Diego M Carrion, Donald Nyanhongo, Edward Bass, Edward Mains, Edwin Chau, Elba Canelon Castillo, Elizabeth Day, Elsayed Desouky, Emily Gaines, Emma Papworth, Emrah Yuruk, Enes Kilic, Eoin Dinneen, Erika Palagonia, Evangelos Xylinas, Faizan Khawaja, Clíodhna Browne, Conor McCann, Cristina Plaza Alonso, David Mak, David Martin-Way, Denula Suthaharan, Dhruv Patel, Diego M Carrion, Donald Nyanhongo, Edward Bass, Edward Mains, Edwin Chau, Elba Canelon Castillo, Elizabeth Day, Elsayed Desouky, Emily Gaines, Emma Papworth, Emrah Yuruk, Enes Kilic, Eoin Dinneen, Erika Palagonia, Evangelos Xylinas, Faizan Khawaja, Fernando Cimarra, Florian Bardet, Francesca Kum, Francesca Peters, Gábor Kovács, Geroge Tanasescu, Giles Hellawell, Giovanni Tasso, Gitte Lam, Giuseppe La Montagna, Giuseppe Pizzuto, Gordan Lenart, Graeme MacLennan, Günal Özgür, Hai Bi, Hannah Lyons, Hannah Warren, Hashim Ahmed, Helen Simpson, Helena Burden, Helena Gresty, Hernado Rios Pita, Holly Clarke, Hosam Serag, Howard Kynaston, Hugh Crawford-Smith, Hugh Mostafid, Hugo Otaola-Arca, Hui Fen Koo, Ibrahim Ibrahim, Idir Ouzaïd, Ignacio Puche-Sanz, Igor Tomašković, Ilker Tinay, Iqbal Sahibzada, Isaac Thangasamy, Iván Revelo Cadena, Jacques Irani, Jakub Udzik, James Brittain, James Catto, James Green, James Tweedle, Jamie Borrego Hernando, Jamie Leask, Jas Kalsi, Jason Frankel, Jason Toniolo, Jay D. Raman, Jean Courcier, Jeevan Kumaradeevan, Jennifer Clark, Jennifer Jones, Jeremy Yuen-Chun Teoh, John Iacovou, John Kelly, John P. Selph, Jonathan Aning, Jon Deeks, Jonathan Cobley, Jonathan Olivier, Jonny Maw, José Antonio Herranz-Yagüe, Jose Ignacio Nolazco, Jose Manuel Cózar-Olmo, Joseph Bagley, Joseph Jelski, Joseph Norris, Joseph Testa, Joshua Meeks, Juan Hernandez, Juan Luis Vásquez, Karen Randhawa, Karishma Dhera, Katarzyna Gronostaj, Kathleen Houlton, Kathleen Lehman, Kathryn Gillams, Kelvin Adasonla, Kevin Brown, Kevin Murtagh, Kiki Mistry, Kim Davenport, Kosuke Kitamura, Laura Derbyshire, Laurence Clarke, Lawrie Morton, Levin Martinez, Louise Goldsmith, Louise Paramore, Luc Cormier, Lucio Dell'Atti, Lucy Simmons, Luis Martinez-Piñeiro, Luis Rico, Luke Chan, Luke Forster, Lulin Ma, Madeline Moore, Maria Camacho Gallego, Maria José Freire, Mark Emberton, Mark Feneley, Marta Antón-Juanilla, Marta Viridiana Muñoz Rivero, Matea Pirša, Matteo Tallè, Matthew Crockett, Matthew Liew, Matthew Trail, Max Peters, Meghan Cooper, Meghana Kulkarni, Michael Ager, Ming He, Mo Li, Mohamed Omran Breish, Mohamed Tarin, Mohammed Aldiwani, Mudit Matanhelia, Muhammad Pasha, Mustafa Kaan Akalin, Nasreen Abdullah, Nathan Hale, Neha Gadiyar, Neil Kocher, Nicholas Bullock, Nicholas Campaign, Nicola Pavan, Nihad Al-Ibraheem, Nikita Bhatt, Nishant Bedi, Nitin Shrotri, Niyati Lobo, Olga Balderas, Omar Kouli, Otakar Capoun, Pablo Oteo Manjavacas, Paolo Gontero, Paramanathan Mariappan, Patricio Garcia Marchiñena, Paul Erotocritou, Paul Sweeney, Paula Planelles, Peter Acher, Peter C. Black, Peter K Osei-Bonsu, Peter Østergren, Peter Smith, Peter-Paul Michiel Willemsse, Piotr L. Chlosta, Qurat Ul Ain, Rachel Barratt, Rachel Esler, Raihan Khalid, Ray Hsu, Remigiusz Stamirowski, Reshma Mangat, Ricardo Cruz, Ricky Ellis, Robert Adams, Robert Hessel, Robert J.A. Oomen, Robert McConkey, Robert Ritchie, Roberto Jarimba, Rohit Chahal, Rosado Mario Andres, Rosalyn Hawkins, Rotimi David, Rustom P. Manecksha, Sachin Agrawal, Syed Sami Hamid, Samuel Deem, Sanchia Goonewardene, Satchi Kuchibhotla Swami, Satoshi Hori, Shahid Khan, Shakeel Mohammad Inder, Shanthi Sangaralingam, Shekhar Marathe, Sheliyan Raveenthiran, Shigeo Horie, Shomik Sengupta, Sian Parson, Sidney Parker, Simon Hawlina, Simon Williams, Simone Mazzoli, Sławomir Grzegorz Kata, Sofia Pinheiro Lopes, Sónia Ramos, Sonpreet Rai, Sophie Rintoul-Hoad, Sorcha O'Meara, Steve Morris, Stacey Turner, Stefano Venturini, Stephanos Almpanis, Steven Joniau, Sunjay Jain, Susan Mallett, Sven Nikles, Shahzad, Sylvia Yan, Taeweon Lee, Taha Uçar, Tamsin Drake, Tarq Toma, Teresa Cabañuz Plo, Thierry Bonnin, Tim Muilwijk, Tim Wollin, Timothy Shun Man CHU, Timson Appanna, Tom Brophy, Tom Ellul, Tomas Austin, Tomaž Smrkolj, Tracey Rowe, Troy Sukhu, Trushar Patel, Tullika Garg, Turhan Çaşkurlu, Uros Bele, Usman Haroon, Víctor Crespo-Afín, Víctor Parejo Cortes, Victoria Capapé Poves, Vincent Gnanapragasam, Vineet Gauhar, Vinnie During, Vivek Kumar, Vojtech Fiala, Wasim Mahmalji, Wayne Lam, Yew Fung Chin, Yigit Filtekin, Yih Chyn Phan, Youssed Ibrahim, Zachary A Glaser, Zainal Adwin Abiddin, Zijian Qin, Zsuzsanna Zotter, Zulkifli Zainuddin.

## 1. Introduction

Haematuria is the most common presentation of suspected urinary tract cancers and is the leading cause of referral to secondary care amongst the urological cancer pathways [1,2]. Bladder cancer is the most prevalent cancer diagnosis in these patients, whilst upper tract urothelial cancers (UTUCs) and renal cell carcinomas (RCCs) are rarer [3,4]. Urinary tract cancers carry high morbidity and mortality [5,6], and early diagnosis is important. The investigation of haematuria poses a huge global health burden as there are many causes, including benign pathology [7]. Reducing unnecessary investigations whilst identifying clinically important urinary tract cancers is an important unmet need [8]. As a result, risk markers should be used to select patients for referral and to determine urgency for further investigation. A recent prediction model to identify bladder cancer in patients with haematuria was developed to recommend a threshold for investigation [9]. Four well-established risk markers were used to predict bladder cancer (type of haematuria, age, sex, and smoking history) [7,10–13]. However, the model neither predicted upper tract cancer nor included other important clinical risk markers such as anticoagulation, previous radiotherapy, and urinary tract infections, which have been described in the literature [14–16].

The IDENTIFY study is the largest prospective cohort study of patients referred with suspected urinary tract cancer. One aim of the study was to develop a model to predict urinary tract cancer in patients referred with haematuria, using several predetermined clinical risk markers. The effects of such a wide range of potential risk markers have not been investigated simultaneously in a study of this scale. This is necessary to account for the complex interplay of factors related to an individual's particular risk of cancer. Individualised risk predictions can then guide a shared decision-making process between the patient and the clinician about further investigation.

## 2. Patients and methods

### 2.1. Study design and source of data

The IDENTIFY study was an international prospective cohort study in patients referred to secondary care with suspected urinary tract cancer [17]. Further details of the study including the adjusted prevalence of cancer have been published [4]. Data collected included the reason for referral, baseline demographic information, clinical history, urine analysis, cytology, imaging findings, cystoscopy findings, histopathology from biopsies or surgery, and multidisciplinary team decisions.

In this analysis, we used the cohort of patients with haematuria (visible or nonvisible haematuria) to develop the prediction model. Nonvisible haematuria was defined as a trace or more on urinalysis, or three or more red blood cells per high power field on microscopy. Microscopy was not required to confirm a urinalysis positive for blood. Patient data were obtained from hospital records of consecutive patients attending a secondary care “haematuria clinic” for a flexible cystoscopy between December 2017 and December 2018. Patients were followed up until their haematuria investigations were concluded and a diagnosis was confirmed or ruled out, as per the judgement of the clinical care team.

We report this study according to the Transparent Reporting of a Multi-variable Prediction Model for Individual Prognosis or Diagnosis (TRIPOD) statement (Supplementary material) [18].

### 2.2. Participants

We included patients aged  $\geq 16$  yr, with haematuria, who were referred to a urologist and underwent investigation for suspected urinary tract cancer. The investigations performed were determined by the urologist's standard practice. Patients were excluded if they had a previous or known diagnosis of primary urological cancer or were undergoing investigations for the recurrence of a primary urological cancer.

### 2.3. Outcome

The primary outcome was the absence or presence of any urinary tract cancer (defined as bladder cancer, UTUC, or RCC) [4,17]. We defined cancer as per Supplementary Table 1 [19,20]. We developed the main model to predict urinary tract cancer. This was derived from three secondary prediction models, one for each type of urinary tract cancer. The secondary outcomes were the absence or presence of each type of urinary tract cancer in their respective models.

### 2.4. Candidate predictors

Candidate predictors for each type of cancer were chosen based on clinical reasoning and evidence from literature. Table 1 shows the candidate predictors that were used to develop the initial models for each type of cancer. These were chosen *a priori* as clinically relevant predictors of urinary tract cancer, by the research steering committee and clinical experts during study design [17].

### 2.5. Sample size

We used the fixed sample size of 10 282 patients based on the primary goal of the study, which was the estimated prevalence of urinary tract cancers [4,17].

### 2.6. Statistical analysis methods

We used a multilevel mixed-effect logistic regression model with country and site as random effects, and all other candidate predictors as fixed effects. Age was included as a continuous predictor; all other predictors were categorical. We performed a complete case analysis and did not impute the small proportion of missing data (see the Supplementary material).

We developed separate models for each type of cancer (ie, bladder cancer, UTUC, and RCC) first, as different predictors may be relevant for one type of cancer but not the other. An example is flank pain, which would be hypothesised to be more relevant in UTUC and renal cancer than in bladder cancer. These secondary predictive models for each type of cancer were developed using a combination of clinical judgement and the backward stepwise elimination process for the poorly understood predictors. We initially fitted the full multivariable model and then performed backward stepwise elimination as we were exploring new candidate predictors. Crucially though, we judged the clinical importance of keeping each predictor before it was eliminated, as clinical selection of predictors is more important than statistical methods alone. The performance of each model was reported.

We fitted statistically significant two-way interaction terms that we deemed clinically important and made clinical sense. The rationale for fitting interaction terms was to account for differences in the predictor's effect between subgroups. A *p* value of  $<0.05$  was considered statistically significant.

**Table 1 – List of predictors for bladder cancer, upper tract urothelial cancer (UTUC), and renal cancer**

| <i>Predictors for all cancers</i>   |                                     |                                |
|---|-------------------------------------|--------------------------------|
| Type of haematuria (visible or nonvisible haematuria)   |                                     |                                |
| Age in years  |                                     |                                |
| Sex (male, female)  |                                     |                                |
| Smoking history (current, ex-smoker, never smoker, or unknown)  |                                     |                                |
| High-risk occupation (yes/no or unknown), defined as exposure to occupational hazards associated with bladder cancer, for example, dyes, rubber, textiles, and pesticides |                                     |                                |
| High-risk travel (yes/no or unknown)—risk of schistosomiasis in freshwater lakes in Africa, South America, or Middle East   |                                     |                                |
| High-risk medications, for example, cyclophosphamide and pioglitazone (yes/no or unknown)   |                                     |                                |
| Episode of urinary tract infection associated with the patient's presentation (none/single/recurrent)   |                                     |                                |
| Anticoagulation (yes/no)—includes warfarin, novel anticoagulant (eg, rivaroxaban, apixaban), antiplatelet (eg, aspirin, clopidogrel), and heparin (any)                   |                                     |                                |
| Previous negative investigation for haematuria (yes/no)   |                                     |                                |
| <i>Predictors specific to type of cancer</i>  |                                     |                                |
| Bladder cancer  | UTUC                                | Renal cancer                   |
| Family history of urothelial cancer (yes/no)  | Family history of urothelial cancer | Family history of renal cancer |
| Dysuria or suprapubic pain (yes/no)   | Flank pain (yes/no)                 | Flank pain (yes/no)            |
| Any lower urinary tract symptoms (obstructive/voiding, storage/irritative or mixed)   |                                     |                                |
| The current use of a catheter (yes/no), including urethral, suprapubic, and intermittent  |                                     |                                |
| Ethnicity (White, Asian, Black/African American, other)   |                                     |                                |
| History of previous pelvic radiotherapy (yes/no)  |                                     |                                |

The predictors used in the predictive model for all urinary tract cancers were selected from all three secondary models, based on the clinical judgement of the study's steering committee. Further detailed statistical analysis methods are included in the [Supplementary material](#).

### 2.7. Evaluation of performance

The calibration slope, area under the receiver operating characteristic curve (AUC), and decision curve analysis were used to assess the performance of the final predictive model for all urinary tract cancers. The model was internally validated using bootstrap resampling with 200 repetitions, and the optimism was calculated, as per the TRIPOD statement (ie, development of the model in each bootstrap sample) [18]. Furthermore, we evaluated the performance (AUC) of the model in different countries. We chose not to split the data into a development and validation cohort firstly due to the low number of events in the rarer cancers, which would limit the number of candidate predictors in their models, and secondly as we intend to perform a separate study to externally validate the prediction model recommended by the TRIPOD statement.

### 2.8. Risk calculation tool

The risk score was created using the coefficients from the linear equation of the multivariable logistic regression. We subsequently developed an online calculator rather than a nomogram to predict an individual's probability of urinary tract cancer, for its ease of use and to demonstrate the tool.

### 2.9. Development of risk score and risk groups

As a guide, we developed risk groups to aid clinicians in determining further investigation. To this end, we defined four categories of cancer risk: very low, low, intermediate, and high. The justification for a very low risk category was to identify a group where investigation may be avoided or delayed. Low-, intermediate-, and high-risk groups are commonly used in clinical practice and reflect a gradient where increasing intensity and urgency of investigation are required. We selected a threshold of <1% predicted risk for the very low group, and 5% and 20% as cut-offs to create low-, intermediate-, and high-risk groups. All thresholds were selected based on clinical reasoning by the study steering committee, as there are no established risk stratification thresholds in use within secondary care in this field. Clinical reasoning was felt to be more appropriate and meaningful than statistical methods for risk thresholds. The

observed cancer prevalence from the cohort was cross-checked in each risk group to ensure that they reflected clinically appropriate stratification.

All analyses were performed using Stata version 16.1 (StataCorp, College Station, TX, USA). This study was registered with clinicaltrials.gov NCT03548688, and the study protocol was published in advance [17].

## 3. Results

### 3.1. Participants

[Figure 1](#) describes the flow of patients through the study. The prevalence was 17.2% ( $n = 1763$ ) for bladder cancer, 1.20% ( $n = 123$ ) for UTUC, and 1.00% ( $n = 103$ ) for renal cancer. The clinical characteristics of the cohort are described in [Table 2](#) and include patient demographics and all candidate predictors.

### 3.2. Model development and specification

Supplementary Tables 2–7 show model development for each type of cancer (candidate predictors and number of patients included in each backward stepwise elimination), in addition to the coefficients and odds ratios of variables in their respective final prediction models. The final predictors in the UTUC model were type of haematuria, age, flank pain, and smoking history, and those in the RCC model were type of haematuria and flank pain. There were 11 final predictors in the bladder cancer model. [Table 3](#) shows the main predictive model for urinary tract cancers using selected predictors from all three secondary models. The number of events (cancers) in the main model was 1863 (19.7%).

The AUC of the final model was 0.86 (95% confidence interval [CI] 0.85–0.87), its calibration slope was 1.03 (95% CI 0.98–1.09), and its intercept was –0.01 (95% CI –0.02 to 0.01) [21]. The decision curve analysis shows a net benefit over investigating all or none ([Supplementary Fig. 1](#)). The optimism between the bootstrap model and the test model was 0.005 (difference between an apparent model AUC of 0.858 and a corrected model AUC of 0.853). Evaluation of

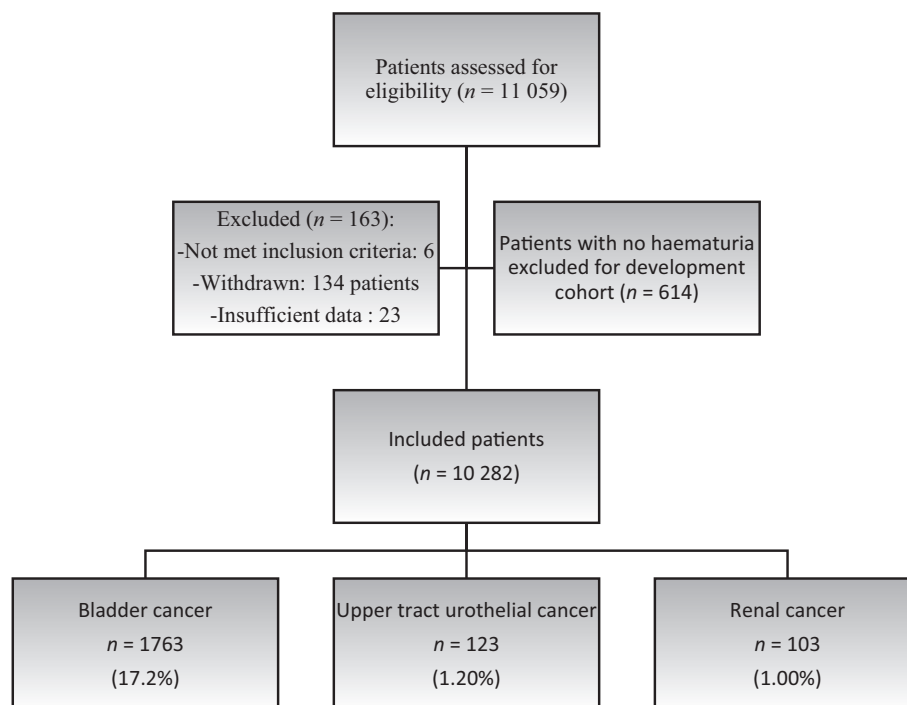


Fig. 1 – Cohort flow diagram.

Table 2 – Patient demographics and clinical characteristics

|  | Total<br>n (%) | No cancer<br>n (%) | Any urinary tract cancer<br>n (%) | Bladder cancer<br>n (%) | Upper tract urothelial cancer<br>n (%) | Renal cancer<br>n (%) |
|--|----------------|--------------------|-----------------------------------|-------------------------|--|-----------------------|
| Total  | 10 282         | 8329 (81.0)        | 1953 (19.0)                       | 1763 (17.1)             | 123 (1.20)                             | 103 (1.00)            |
| <b>Type of haematuria</b>                    |                |                    |                                   |                         |  |                       |
| Nonvisible haematuria                        | 3152 (30.7)    | 2967 (35.6)        | 185 (9.47)                        | 165 (9.36)              | 9 (7.32)                               | 13 (12.6)             |
| Visible haematuria                           | 7130 (69.3)    | 5362 (64.4)        | 1768 (90.5)                       | 1598 (90.6)             | 114 (92.7)                             | 90 (87.4)             |
| Age (yr), mean (SD)                          | 64.3 (14.6)    | 62.8 (14.8)        | 70.4 (11.9)                       | 70.7 (11.8)             | 71.9 (11.7)                            | 64.8 (13.1)           |
| <b>Sex</b>                                   |                |                    |                                   |                         |  |                       |
| Female                                       | 3853 (37.5)    | 3384 (40.6)        | 469 (24.0)                        | 413 (23.4)              | 40 (32.5)                              | 25 (24.3)             |
| Male   | 6423 (62.5)    | 4940 (59.3)        | 1483 (75.9)                       | 1349 (76.5)             | 83 (67.5)                              | 78 (75.7)             |
| Other  | 6 (0.06)       | 5 (0.06)           | 1 (0.05)                          | 1 (0.06)                | 0 (0)                                  | 0 (0)                 |
| <b>Smoking</b>                               |                |                    |                                   |                         |  |                       |
| Never smoker                                 | 4651 (45.2)    | 4099 (49.2)        | 552 (28.3)                        | 477 (27.1)              | 41 (33.3)                              | 44 (42.7)             |
| Ex-smoker                                    | 3066 (29.2)    | 2267 (27.2)        | 739 (37.8)                        | 680 (38.6)              | 38 (30.9)                              | 36 (35.0)             |
| Current smoker                               | 1881 (18.3)    | 1367 (16.4)        | 514 (26.3)                        | 471 (26.7)              | 35 (28.5)                              | 17 (16.5)             |
| Unknown                                      | 744 (7.24)     | 596 (7.16)         | 148 (7.58)                        | 135 (7.66)              | 9 (7.32)                               | 6 (5.83)              |
| <b>Occupational risk<sup>a</sup></b>         |                |                    |                                   |                         |  |                       |
| No   | 8548 (83.1)    | 6958 (83.5)        | 1590 (81.4)                       | 1428 (81.0)             | 102 (82.9)                             | 90 (87.4)             |
| Yes  | 393 (3.82)     | 279 (3.35)         | 114 (5.84)                        | 109 (6.18)              | 5 (4.07)                               | 2 (1.94)              |
| Unknown                                      | 993 (9.66)     | 784 (9.41)         | 209 (10.7)                        | 190 (10.8)              | 13 (10.6)                              | 9 (8.74)              |
| Missing                                      | 348 (3.38)     | 308 (3.70)         | 40 (2.05)                         | 36 (2.04)               | 3 (2.44)                               | 2 (1.94)              |
| <b>Medication risk<sup>b</sup></b>           |                |                    |                                   |                         |  |                       |
| No   | 9203 (89.5)    | 7459 (89.6)        | 1744 (89.3)                       | 1575 (89.3)             | 108 (87.8)                             | 93 (90.3)             |
| Yes  | 77 (0.75)      | 57 (0.68)          | 20 (1.02)                         | 17 (0.96)               | 2 (1.63)                               | 1 (0.97)              |
| Unknown                                      | 628 (6.11)     | 478 (5.74)         | 150 (7.68)                        | 137 (7.77)              | 9 (7.32)                               | 7 (6.80)              |
| Missing                                      | 374 (3.64)     | 335 (4.02)         | 39 (2.00)                         | 34 (1.93)               | 4 (3.25)                               | 2 (1.94)              |
| <b>Travel/environmental risk<sup>c</sup></b> |                |                    |                                   |                         |  |                       |
| No   | 8761 (85.2)    | 7030 (84.4)        | 1731 (88.6)                       | 1566 (88.8)             | 105 (85.4)                             | 91 (88.4)             |
| Yes  | 108 (1.05)     | 99 (1.19)          | 9 (0.46)                          | 9 (0.51)                | 0 (0)                                  | 0 (0)                 |
| Unknown                                      | 948 (9.22)     | 786 (9.44)         | 162 (8.29)                        | 143 (8.11)              | 13 (10.6)                              | 10 (9.71)             |
| Missing                                      | 465 (4.52)     | 414 (4.97)         | 51 (2.61)                         | 45 (2.55)               | 5 (4.07)                               | 2 (1.94)              |
| <b>UTI history</b>                           |                |                    |                                   |                         |  |                       |
| None   | 7873 (76.6)    | 6155 (73.9)        | 1718 (88.0)                       | 1548 (87.8)             | 110 (89.4)                             | 92 (89.3)             |
| Single                                       | 1250 (12.2)    | 1123 (13.5)        | 127 (6.50)                        | 115 (6.52)              | 8 (6.50)                               | 6 (5.83)              |
| Recurrent                                    | 1018 (9.90)    | 930 (11.2)         | 88 (4.51)                         | 80 (4.54)               | 5 (4.07)                               | 5 (4.85)              |
| Missing                                      | 141 (1.37)     | 121 (1.45)         | 20 (1.02)                         | 20 (1.13)               | 0 (0)                                  | 0 (0)                 |
| <b>Dysuria/suprapubic pain</b>               |                |                    |                                   |                         |  |                       |
| No   | 7909 (76.9)    | 6300 (75.6)        | 1609 (82.4)                       | 1442 (81.8)             | 111 (90.2)                             | 84 (81.6)             |

(continued on next page)



Table 2 (continued)

|                                     | Total<br>n (%) | No cancer<br>n (%) | Any urinary tract cancer<br>n (%) | Bladder cancer<br>n (%) | Upper tract urothelial cancer<br>n (%) | Renal cancer<br>n (%) |
|-------------------------------------|----------------|--------------------|-----------------------------------|-------------------------|--|-----------------------|
| Yes                                 | 2144 (20.9)    | 1830 (22.0)        | 314 (16.1)                        | 292 (16.6)              | 11 (8.94)                              | 19 (18.4)             |
| Missing                             | 229 (2.23)     | 199 (2.39)         | 30 (1.54)                         | 29 (1.64)               | 1 (0.81)                               | 0 (0)                 |
| Previous haematuria evaluation      |                |                    |                                   |                         |  |                       |
| No                                  | 9130 (88.8)    | 7312 (87.8)        | 1818 (93.1)                       | 1640 (93.0)             | 115 (93.5)                             | 96 (93.2)             |
| Yes                                 | 1023 (9.95)    | 907 (10.9)         | 116 (5.94)                        | 104 (5.90)              | 8 (6.50)                               | 7 (6.80)              |
| Missing                             | 129 (1.25)     | 110 (1.32)         | 19 (0.97)                         | 19 (1.08)               | 0 (0)                                  | 0 (0)                 |
| Family history of urothelial cancer |                |                    |                                   |                         |  |                       |
| No                                  | 9734 (94.7)    | 7874 (94.5)        | 1860 (95.2)                       | 1680 (95.3)             | 119 (96.8)                             | 96 (93.2)             |
| Yes                                 | 219 (2.13)     | 157 (1.88)         | 62 (3.17)                         | 55 (3.12)               | 2 (1.63)                               | 5 (4.85)              |
| Missing                             | 329 (3.20)     | 298 (3.58)         | 31 (1.59)                         | 28 (1.59)               | 2 (1.63)                               | 2 (1.94)              |
| LUTS                                |                |                    |                                   |                         |  |                       |
| None                                | 5991 (58.3)    | 4765 (57.2)        | 1226 (62.8)                       | 1096 (62.2)             | 82 (66.7)                              | 67 (65.1)             |
| Obstructive/voiding                 | 1354 (13.2)    | 1085 (13.0)        | 269 (13.8)                        | 247 (14.0)              | 14 (11.4)                              | 13 (12.6)             |
| Storage/irritative                  | 1800 (17.5)    | 1543 (18.5)        | 257 (13.2)                        | 232 (13.2)              | 15 (12.2)                              | 15 (14.6)             |
| Mixed                               | 1010 (9.82)    | 826 (9.92)         | 184 (9.42)                        | 171 (9.70)              | 12 (9.76)                              | 8 (7.77)              |
| Missing                             | 127 (1.24)     | 110 (1.32)         | 17 (0.87)                         | 17 (0.96)               | 0 (0)                                  | 0 (0)                 |
| Use of catheter                     |                |                    |                                   |                         |  |                       |
| No                                  | 9833 (95.6)    | 7930 (95.2)        | 1903 (97.4)                       | 1715 (97.3)             | 120 (97.6)                             | 101 (98.1)            |
| Yes                                 | 334 (3.25)     | 301 (3.61)         | 33 (1.69)                         | 31 (1.76)               | 3 (2.44)                               | 2 (1.94)              |
| Missing                             | 115 (1.12)     | 98 (1.18)          | 17 (0.87)                         | 17 (0.96)               | 0 (0)                                  | 0 (0)                 |
| Ethnicity                           |                |                    |                                   |                         |  |                       |
| White                               | 7930 (77.1)    | 6288 (75.5)        | 1642 (84.1)                       | 1482 (84.1)             | 108 (87.8)                             | 85 (82.5)             |
| Asian                               | 1195 (11.6)    | 1014 (12.2)        | 181 (9.27)                        | 170 (9.64)              | 6 (4.88)                               | 7 (6.80)              |
| Black                               | 297 (2.89)     | 277 (3.33)         | 20 (1.02)                         | 14 (0.79)               | 3 (2.44)                               | 3 (2.91)              |
| Other                               | 518 (5.04)     | 450 (5.40)         | 68 (3.48)                         | 60 (3.40)               | 3 (2.44)                               | 5 (4.85)              |
| Missing                             | 342 (3.33)     | 300 (3.60)         | 42 (2.15)                         | 37 (2.10)               | 3 (2.44)                               | 3 (2.91)              |
| Pelvic radiotherapy                 |                |                    |                                   |                         |  |                       |
| No                                  | 10 043 (97.7)  | 8125 (97.6)        | 1918 (98.2)                       | 1731 (98.2)             | 123 (100)                              | 100 (97.1)            |
| Yes                                 | 203 (1.97)     | 172 (2.07)         | 31 (1.59)                         | 28 (1.59)               | 0 (0)                                  | 3 (2.91)              |
| Missing                             | 36 (0.35)      | 32 (0.38)          | 4 (0.20)                          | 4 (0.23)                | 0 (0)                                  | 0 (0)                 |
| Anticoagulation                     |                |                    |                                   |                         |  |                       |
| No                                  | 7306 (71.1)    | 6027 (72.4)        | 1279 (65.5)                       | 1153 (65.4)             | 83 (67.5)                              | 69 (67.0)             |
| Yes                                 | 2611 (25.4)    | 1981 (23.8)        | 630 (32.3)                        | 569 (32.3)              | 38 (30.9)                              | 33 (32.0)             |
| Missing                             | 365 (3.55)     | 321 (3.85)         | 44 (2.25)                         | 41 (2.33)               | 2 (1.63)                               | 1 (0.97)              |
| Flank pain                          |                |                    |                                   |                         |  |                       |
| No                                  | 9131 (88.8)    | 7310 (87.8)        | 1821 (93.2)                       | 1670 (94.7)             | 97 (78.9)                              | 87 (84.5)             |
| Yes                                 | 922 (8.97)     | 820 (9.85)         | 102 (5.22)                        | 64 (3.63)               | 25 (20.3)                              | 16 (15.5)             |
| Missing                             | 229 (2.23)     | 199 (2.39)         | 30 (1.54)                         | 29 (1.64)               | 1 (0.81)                               | 0 (0)                 |

LUTS = lower urinary tract symptom; SD = standard deviation; UTI = urinary tract infection.

Percentages are column percentages except in the first row ("Total"), which are row percentages. Individual cancers do not add up to total cancers as some patients were diagnosed with more than one type of cancer.

Occupational, medication, and travel/environmental risks were considered regardless of time since exposure.

<sup>a</sup> Defined as exposure to dyes, rubber, textiles, and pesticides.

<sup>b</sup> For example, cyclophosphamide and pioglitazone.

<sup>c</sup> Risk of schistosomiasis: relevant exposure to freshwater lakes in Africa, South America, and Middle East.

Table 3 – Final prediction model for urinary tract cancer using mixed effects multivariable logistic regression

| Variable                                 | Coefficient | Odds ratio | 95% Confidence interval | p value |
|--|-------------|------------|-------------------------|---------|
| Nonvisible haematuria                    |             | 1.00       |                         |         |
| Visible haematuria                       | 1.99        | 7.29       | 5.24–10.1               | <0.001  |
| Female                                   |             | 1.00       |                         |         |
| Male                                     | 0.69        | 2.00       | 1.40–2.87               | <0.001  |
| Age (yr)                                 | 0.07        | 1.07       | 1.06–1.09               | <0.001  |
| Age per 5-yr difference                  | 0.15        | 1.17       | 1.10–1.23               | <0.001  |
| Never smoker                             |             | 1.00       |                         |         |
| Ex-smoker                                | 0.70        | 2.02       | 1.74–2.34               | <0.001  |
| Current smoker                           | 1.06        | 2.88       | 2.44–3.41               | <0.001  |
| Family history of urothelial cancer      |             |            |                         |         |
| No                                       |             | 1.00       |                         |         |
| Yes                                      | 0.72        | 2.06       | 1.39–3.03               | 0.001   |
| Previous benign haematuria investigation |             |            |                         |         |
| No                                       |             | 1.00       |                         |         |
| Yes                                      | –0.84       | 0.43       | 0.34–0.55               | <0.001  |
| UTI history                              |             |            |                         |         |
| None                                     |             | 1.00       |                         |         |
| Single                                   | –0.74       | 0.48       | 0.38–0.60               | <0.001  |

**Table 3** (continued)

| Variable                           | Coefficient | Odds ratio | 95% Confidence interval | p value |
|------------------------------------|-------------|------------|-------------------------|---------|
| Recurrent Catheter use             | -0.75       | 0.47       | 0.36–0.62               | <0.001  |
| No                                 |             | 1.00       |                         |         |
| Yes                                | -1.57       | 0.21       | 0.14–0.31               | <0.001  |
| Pelvic radiotherapy history        |             |            |                         |         |
| No                                 |             | 1.00       |                         |         |
| Yes                                | -0.59       | 0.56       | 0.35–0.88               | 0.013   |
| Anticoagulation                    |             |            |                         |         |
| No                                 |             | 1.00       |                         |         |
| Yes                                | -0.17       | 0.84       | 0.70–1.01               | 0.060   |
| Dysuria/suprapubic pain            |             |            |                         |         |
| No                                 |             | 1.00       |                         |         |
| Yes                                | -0.32       | 0.72       | 0.61–0.86               | <0.001  |
| Interaction terms                  |             |            |                         |         |
| Visible haematuria & male          | -0.82       | 0.44       | 0.30–0.65               | <0.001  |
| Visible haematuria & age           | -0.02       | 0.98       | 0.96–0.99               | 0.007   |
| Age & anticoagulation              | -0.02       | 0.98       | 0.97–1.00               | 0.007   |
| Intercept                          | -2.79       |            |                         |         |
| Intercountry variance              | 0.84        |            | 0.40–1.78               |         |
| Intercentre variance               | 0.35        |            | 0.25–0.56               |         |
| Intraclass correlation for country | 0.19        |            | 0.10–0.33               |         |
| Intraclass correlation for centre  | 0.26        |            | 0.17–0.38               |         |

AUC = area under the curve for receiver operating characteristics; CI = confidence interval; UTI = Urinary tract infection. Number of observations in model = 9464 (92.0% of cohort); missing data = 818 (8.0%); number of events = 1863/9464 (19.7%); number of country groups = 26 with a mean of 364 observations per group (minimum = 30, maximum = 4294); number of centre groups = 110 with a mean of 85.3 observations per group (minimum = 36, maximum = 611). Age has been centred about its mean. Performance in predicting all urinary tract cancers in cohort: AUC = 0.86 (95% CI 0.85–0.87).

**Table 4 – Validation of the predictive model for urinary tract cancer: AUC of different countries**

| Country  | AUC  |
|----------|------|
| UK       | 0.80 |
| France   | 0.78 |
| Italy    | 0.75 |
| Spain    | 0.75 |
| USA      | 0.84 |
| Canada   | 0.77 |
| Ireland  | 0.85 |
| Portugal | 0.81 |
| Turkey   | 0.78 |
| China    | 0.89 |

AUC = area under the curve for receiver operating characteristics.

the model on different countries showed good performance, with an AUC of at least 0.75 (Italy and Spain) to 0.89 (China; Table 4).

**3.3. Risk calculator**

The predicted risk was calculated from the risk score, which is derived from the linear coefficients as follows:

$$\text{Risk score} = \text{Intercept} + 0.07 * (\text{age} - \text{mean age}) + 0.69 * (\text{male}) + 1.99 * (\text{visible haematuria}) + 1.06 * (\text{smoker}) + 0.70 * (\text{ex-smoker}) + 0.72 * (\text{family history of urothelial cancer}) - 0.84 * (\text{previous benign haematuria investigation}) - 0.74 * (\text{single episode of UTI}) - 0.75 * (\text{recurrent episodes of UTI}) - 1.57 * (\text{catheter use}) - 0.59 * (\text{pelvic radiotherapy history}) - 0.17 * (\text{anticoagulation}) - 0.32 * (\text{dysuria}) - 0.82 * (\text{sex} * \text{visible haematuria}) - 0.02 * (\text{visible haematuria} * [\text{age} - \text{mean age}]) - 0.02 * (\text{anticoagulation} * [\text{age} - \text{mean age}])$$

which relates to an individual's probability of urinary tract cancer as follows:

Patient's individual risk of urinary tract cancer =  $1/1 + \text{exponential}^{-\text{(Risk Score)}}$

where age is a continuous variable in years, and all other variables are assigned a value of 0 if absent and a value of 1 if present.

As an example of its use, a 70-yr-old man with visible haematuria who is a current smoker has a predicted risk of 51.7% and would be classified to have a high risk. Conversely, a 40-yr-old woman with a single urinary tract infection (UTI) associated with visible haematuria has a predicted risk of 4.0% and would be classified to have a low risk.

When stratified by risk groups, the majority of patients (over 80%) from our cohort were stratified into intermediate- and high-risk groups (Table 5), which also had the largest proportion of cancers (Supplementary Fig. 2). Within the very-low-, low-, intermediate-, and high-risk groups from our cohort, the cancer prevalence was 0.82%, 3.90%, 10.5%, and 30.5%, respectively.

**4. Discussion**

Our principal finding in this analysis was the development of a clinically relevant and practical prediction model for urinary tract cancer, with good discrimination, which can support clinicians in prioritising the investigation of patients with haematuria. This is the first model developed using a broad international cohort and was designed to investigate a number of clinically important risk markers commonly proposed to be associated with cancer detection. The final predictors of increased risk of cancer in our model were visible haematuria, older age, current or ex-smoker history, family history of urothelial cancer, and male sex. Predictors associated with a decreased risk of cancer in

**Table 5 – Stratification of observed cancers by risk categories**

|  | Very low risk | Low risk    | Intermediate risk | High risk   | Total             |
|--|---------------|-------------|-------------------|-------------|-------------------|
| Number of patients, <i>n</i> (row %)   | 366 (3.56)    | 1411 (13.7) | 3318 (32.3)       | 5187 (50.5) | 10 282            |
| Cancer prevalence, <i>n</i> (column %) | 3 (0.82)      | 55 (3.90)   | 349 (10.5)        | 1581 (30.5) | 1988 <sup>a</sup> |
| Bladder cancer, <i>n</i>               | 2             | 45          | 314               | 1401        | 1762              |
| UTUC cancer, <i>n</i>                  | 0             | 2           | 11                | 110         | 123               |
| Renal cancer, <i>n</i>                 | 1             | 8           | 24                | 70          | 103               |

UTUC = upper tract urothelial cancer.

<sup>a</sup> Some patients had more than one type of cancer, so the total sum is higher than the number of patients with cancer.

the model were previous benign haematuria investigations, UTIs associated with the haematuria presentation, dysuria or suprapubic pain, anticoagulation, catheter use, and previous pelvic radiotherapy.

The study was specifically designed to include a wide variety of risk markers associated with urinary tract cancer in patients with haematuria. Another strength of the study includes its large cohort size (with a large number of events per predictor), which is important for a multivariable analysis especially with rarer cancers. The relatively higher cancer prevalence than that reported in previous studies has been explained further in the prevalence analysis and is due to adjustment of confounders and a more representative, larger, international population sample [4]. Furthermore, results are generalisable within secondary care given the diversity of the cohort, as it is the first multinational study on patients with haematuria. The adjustment for geographical effects within the model takes into consideration the heterogeneity in the background risk of different countries.

Our model improves on an existing predictive model [9] by including upper tract cancers and a wider variety of markers, whilst being consistent with the association of visible haematuria, male sex, age, and smoking with bladder cancer. Another difference between models is the use of a risk threshold for investigation versus risk stratification of all referred patients. By using a risk threshold over which patients should be investigated, cancers may be missed if their predicted risk is below this threshold. By applying it to our cohort, 53/1953 (2.71%) of all urinary tract cancers would be missed. Conversely, our model uses risk stratification and so considers all referred patients for investigation.

Some risk markers in our model showed a decreased risk of cancer, likely because their presence is a sign of benign causes of haematuria rather than malignant disease. In patients with UTIs and haematuria, prior studies did not find a difference in the cancer risk between these patients and a control group [7,16]. The association of pelvic radiotherapy with bladder cancer has mainly been shown in selected patients with prostate cancer comparing radiotherapy with radical surgery [14,22], not in patients presenting with haematuria. Similarly, evidence showing the association of catheter use with bladder cancer is limited and demonstrated in a selected population without haematuria [23]. Our study population is therefore different, and our analysis suggests that these markers are more likely associated with benign disease such as cystitis [24].

It has been reported that the most common cause for haematuria admissions to hospital among patients on an oral anticoagulant is benign disease (21%), followed by

urothelial carcinoma (17%) [15]. This may explain its effect in our model as it was generally associated with a reduced risk of cancer, except in older patients due to its interaction with age. We included anticoagulation as a predictor in the final model even though it was not statistically significant for urinary tract cancers ( $p = 0.06$ ), due to its clinical importance and common use in patients with haematuria.

The main limitation of the prediction model is that it excludes patients with a prior history of urological malignancy or without haematuria, and it should not be applied to such patients. Furthermore, one significant predictor (flank pain) was excluded from the final model for urinary tract cancer following clinical judgement, as upper tract cancers are much rarer than bladder cancer in patients with haematuria. In addition, flank pain is commonly associated with benign pathology such as urolithiasis. However, we suggest that the model may be modified to upstage patients' risk category if they had flank pain and haematuria. Secondly, although we conducted thorough internal validation, we were unable to perform external validation on a separate dataset as, to our knowledge, such an extensive dataset of predictors in patients referred to secondary care has not been published or made available.

After external validation, we envisage the typical use of our predictive model, as an online calculator, by urologists in a "haematuria clinic" setting to prioritise and triage patients referred to them with any type of haematuria based on the overall risk of urinary tract cancer. We would also recommend this tool to be used to prompt a shared decision-making process between the clinician and the patient regarding their individual risks and to guide the urgency and necessity of investigations. Though we have set thresholds for risk stratification, clinicians may use their own judgement of risk based on the patients' individual predicted risks of cancer. These thresholds were also the most voted thresholds for use in a Twitter poll by urologists [25]. The threshold of <1% for very-low-risk cancer was chosen as this is less than the risk of infection in patients undergoing a flexible cystoscopy [26]. Therefore, avoidance of investigation in these patients may be considered. Prioritising high-risk patients for early investigation and detection of urinary tract cancer are important for reducing morbidity and mortality, and for prioritising the use of limited resources. Similarly, avoiding or delaying investigation in very-low-risk patients and choosing a less invasive or lower urgency investigative approach in low-risk patients may improve patient experience and resource burden.

Additional research to build on this analysis should include external validation of the model and assessment of its use in a clinical trial evaluating a new diagnostic path-

way for patients referred with haematuria. Further recommendations for the clinical implication and type of investigation required in each risk category, especially with regard to imaging and cytology, require diagnostic test evaluation and will be the subject of future analysis from the IDENTIFY study.

## 5. Conclusions

We present a risk prediction model for the detection of urinary tract cancer based on a large international cohort of patients presenting with haematuria in secondary care. It can be used in secondary care for risk stratification and to aid the shared decision-making process between clinicians and patients for any further investigation. This could have a major impact on healthcare resource usage.

**Author contributions:** Sinan Khadhoury had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

*Study concept and design:* Khadhoury, McGrath, Kasivisvanathan, Shah, Gallagher, MacKenzie.

*Acquisition of data:* All listed authors and collaborators.

*Analysis and interpretation of data:* Khadhoury, MacLennan, Goulao, Gallagher, Kasivisvanathan.

*Drafting of the manuscript:* Khadhoury, Shah, Gallagher, Kasivisvanathan.

*Critical revision of the manuscript for important intellectual content:* Gallagher, MacKenzie, Shah, Gao, Moore, Zimmermann, Edison, Jefferies, Nambiar, Anbarasan, Mannas, Lee, Marra, Rivas, Marcq, Assmus, Uçar, Boltri, La Montagna, Burnhope, Nkwam, Austin, Boxall, Downey, Sukhu, Antón-Juanilla, Rai, Chin, Moore, Drake, Green, Goulao, MacLennan, Nielsen, McGrath, Kasivisvanathan.

*Statistical analysis:* Khadhoury, MacLennan, Goulao.

*Obtaining funding:* Khadhoury, Kasivisvanathan.

*Administrative, technical, or material support:* Khadhoury, MacKenzie, Shah, Gao, Moore, Zimmermann, Edison.

*Supervision:* Kasivisvanathan, McGrath.

*Other:* None.

**Financial disclosures:** Sinan Khadhoury certifies that all conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject matter or materials discussed in the manuscript (eg, employment/affiliation, grants or funding, consultancies, honoraria, stock ownership or options, expert testimony, royalties, or patents filed, received, or pending), are the following: None.

**Funding/Support and role of the sponsor:** Grants from Action Bladder Cancer UK, The Urology Foundation, The Rosetrees Trust were used for costs of statistical analysis and dissemination of results at international meetings and conferences. There were no endorsements from pharmaceutical companies or agencies to write this article. Veeru Kasivisvanathan is an Academic Clinical Lecturer funded by the United Kingdom National Institute for Health Research (NIHR). The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR or the Department of Health and Social Care.

**Acknowledgments:** We would like to thank all the BURST research collaborators for taking part in this study, Max Peters for his support and advice regarding the methods and Jonathan Deeks for his support from the Test

Evaluation Research Group. Though unrelated to this study, the BURST Research Collaborative would like to acknowledge funding from the BJU International, the British Association of Urological Surgeons, Ferring Pharmaceuticals Ltd, and Dominvs Group.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.euf.2022.06.001>.

## References

- [1] Mathew A, Desai K. An audit of urology two-week wait referrals in a large teaching hospital in England. *Ann R Coll Surg Engl* 2009;91:310.
- [2] Hawary AM, Warburton HE, Brough RJ, et al. The “2-week wait” rule for referrals for suspected urological cancers—urgent need for refinement of criteria. *Ann R Coll Surg Engl* 2008;90:517–22.
- [3] Tan WS, Feber A, Sarpong R, et al. Who should be investigated for haematuria? Results of a contemporary prospective observational study of 3556 patients. *Eur Urol* 2018;74:10–4.
- [4] Khadhoury S, Gallagher KM, MacKenzie KR, et al. The IDENTIFY study: the investigation and detection of urological neoplasia in patients referred with suspected urinary tract cancer; a multicentre observational study. *BJU Int* 2021;128:440–50.
- [5] Ferlay J, Colombet M, Soerjomataram I, et al. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer* 2019;144:1941–53.
- [6] Wong MCS, Fung FDH, Leung C, Cheung WWL, Goggins WB, Ng CF. The global epidemiology of bladder cancer: a joinpoint regression analysis of its incidence and mortality trends and projection. *Sci Rep* 2018;8:1129.
- [7] Khadra MH, Pickard RS, Charlton M, Powell PH, Neal DE. A prospective analysis of 1,930 patients with hematuria to evaluate current diagnostic practice. *J Urol* 2000;163:524–7.
- [8] Rodgers MA, Hempel S, Aho T, Kelly JD, Kleijnen J, Westwood M. Diagnostic tests used in the investigation of adult haematuria: a systematic review. *BJU Int* 2006;98:1154–60.
- [9] Tan WS, Ahmad A, Feber A, et al. Development and validation of a haematuria cancer risk score to identify patients at risk of harbouring cancer. *J Intern Med* 2019;285:436–45.
- [10] Edwards TJ, Dickinson AJ, Natale S, Gosling J, McGrath JS. A prospective analysis of the diagnostic yield resulting from the attendance of 4020 patients at a protocol-driven haematuria clinic. *BJU Int* 2006;97:301–5.
- [11] Dobruch J, Daneshmand S, Fisch M, et al. Gender and bladder cancer: a collaborative review of etiology, biology, and outcomes. *Eur Urol* 2016;69:300–10.
- [12] Brennan P, Bogillot O, Cordier S, et al. Cigarette smoking and bladder cancer in men: a pooled analysis of 11 case-control studies. *Int J Cancer* 2000;86:289–94.
- [13] Schmidt-Hansen M, Berendse S, Hamilton W. The association between symptoms and bladder or renal tract cancer in primary care: a systematic review. *Br J Gen Pract* 2015;65:e769–75.
- [14] Moschini M, Zaffuto E, Karakiewicz PI, et al. External beam radiotherapy increases the risk of bladder cancer when compared with radical prostatectomy in patients affected by prostate cancer: a population-based analysis. *Eur Urol* 2019;75:319–28.
- [15] Satasivam P, Reeves F, Lin M, et al. The effect of oral anticoagulation on the prevalence and management of haematuria in a contemporary Australian patient cohort. *BJU Int* 2012;110(Suppl 4):80–4.
- [16] Summerton N, Mann S, Rigby AS, Ashley J, Palmer S, Hetherington JW. Patients with new onset haematuria: assessing the discriminant value of clinical information in relation to urological malignancies. *Br J Gen Pract* 2002;52:284–9.
- [17] Khadhoury S, Gallagher KM, MacKenzie K, et al. IDENTIFY: The investigation and detection of urological neoplasia in patients referred with suspected urinary tract cancer: a multicentre cohort study. *Int J Surg Protoc* 2020;21:8–12.
- [18] Moons KGM, Altman DG, Reitsma JB, et al. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1–W73.

- 
- [19] Moch H, Cubilla AL, Humphrey PA, Reuter VE, Ulbright TM. The 2016 WHO classification of tumours of the urinary system and male genital organs—part A: renal, penile, and testicular tumours. *Eur Urol* 2016;70:93–105.
- [20] Humphrey PA, Moch H, Cubilla AL, Ulbright TM, Reuter VE. The 2016 WHO classification of tumours of the urinary system and male genital organs—part B: prostate and bladder tumours. *Eur Urol* 2016;70:106–19.
- [21] Ensor J, Snell KI, Martin EC. PMCALPLOT: Stata module to produce calibration plot of prediction model performance. *Statistical Software Components* 2020.
- [22] Zhao S, Xie Q, Yang R, et al. High prevalence of secondary bladder cancer in men on radiotherapy for prostate cancer: evidence from a meta-analysis. *Cancer Manag Res* 2019;11:587–98.
- [23] Ho CH, Sung KC, Lim SW, et al. Chronic indwelling urinary catheter increase the risk of bladder cancer, even in patients without spinal cord injury. *Medicine (Baltimore)* 2015;94:e1736.
- [24] Gurka MK, Chen LN, Bhagat A, et al. Hematuria following stereotactic body radiation therapy (SBRT) for clinically localized prostate cancer. *Radiat Oncol* 2015;10:44.
- [25] BJU International. For a patient with haematuria referred to you for suspected urinary tract cancer, what risk thresholds would you use to classify them as very low, low, intermediate or high risk? 2021. <https://TwitterCom/BJUJournal/Status/1453448682746044425>.
- [26] Trail M, Cullen J, Fulton E, et al. Evaluating the safety of performing flexible cystoscopy when urinalysis suggests presence of “infection”: results of a prospective clinical study in 2350 patients. *Eur Urol Open Sci* 2021:3128–36.