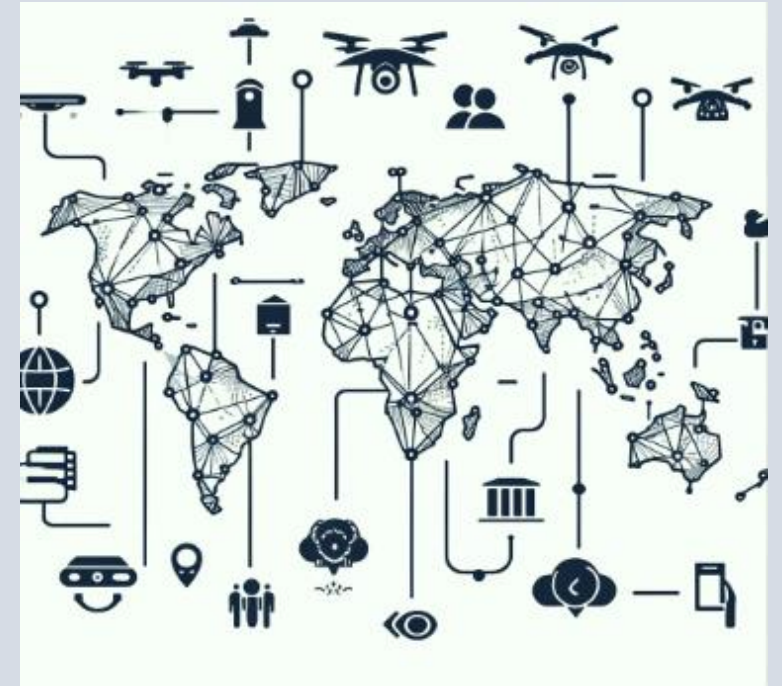


The effectiveness of big data-driven predictive policing: systematic review

Youngsub Lee, Ben Bradford,
Krisztian Posch

Topic: Big-data driven predictive policing

Among the democracies where the use of big data and predictive policing has been established, both the UK and the US have claimed effectiveness as a basis for the legitimacy of implementation in the law enforcement sector (UK House of Parliament, 2014; US Executive Office of the President, 2014; NYC, 2015; NPCC & APCC, 2016).



Concerns over BDPP

At the macro level, the emergence of big data in policing has raised concerns about the 'surveillance society' and the security and transparency issues associated with the collection and retention of big data by police and other state actors (Brayne, 2017; Shapiro, 2017; McGuire, 2021; Zuboff, 2018).



If the data is mishandled during processing, or if it reflects the biases often present in current policing activities and practices, then algorithm-guided policing is also likely to be biased (Brantingham, 2017; Richardson et al., 2019).

Furthermore, the trend towards algorithm-based policing can lead to distrust not only between police and the public, but also within police organisations. (Sandhu & Fussey, 2021; Egbert & Krasmann, 2020; Ratcliffe et al, 2020).



➔ Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence (The White house, 2023)

Scope of this analysis: effectiveness

Despite the concerns raised, the rationale that allows the BDPP to be considered for implementation, namely its '**effectiveness**' in combating crime, should be examined for the first time (claimants side's rationale first).

This means that this review will not deal with the normative or ethical aspects of the BDPP.

It will only look at the instrumental aspects of effectiveness, i.e. whether there has actually been an intended reduction in crime, not at police culture or organisational aspects.

➡ In place-based context*



*In our understanding, place-based policing is the optimisation of police resources based on geospatial data included analysis

Search Terms

-Terms related to big data: "big data" OR "data sources" OR dataset* OR "large data" OR "multiple data" OR "multiple sources" OR "high volume" OR "large volume" OR "high velocity"

AND

-Terms related to policing: crim* OR secur* OR policing OR police OR "law enforcement" OR policed

AND

-Terms related to predictive methods: algorithm* OR "AI" OR "artificial intelligence" OR forecast* OR predict* OR "machine learning" OR "super-learner" OR "super learner"

AND

-Terms related to place-based context: place OR "hot spot" OR spot OR spatial* OR area OR geo*

OR

-Terms related to the actual commercial programmes including "COMPSTAT" OR "Hunchlab" OR "Predpol"

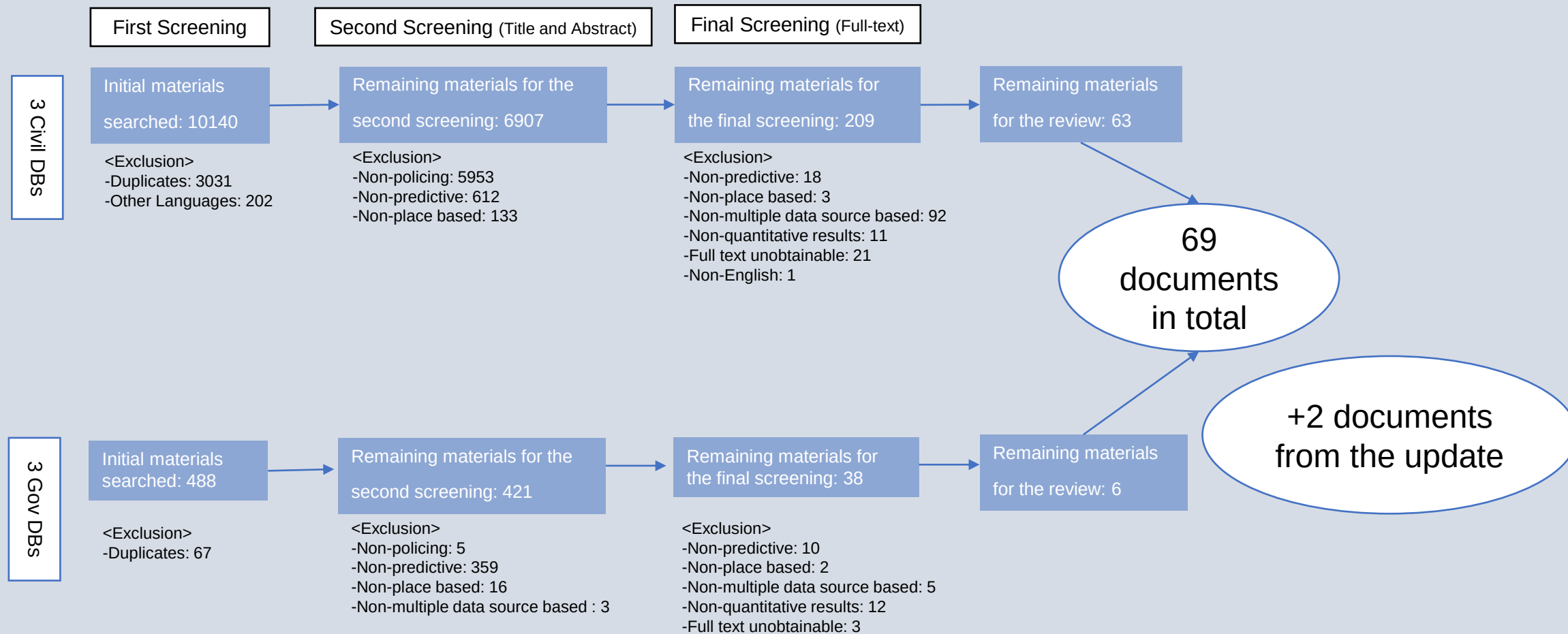
Inclusion & Exclusion Criteria

Criteria		Inclusion	Exclusion
PICO	Population	-Policing strategy using big data(big data: multiple data sources including crime data)	-Not related to the work of police -Only relying on the conventional crime data of police
	Intervention	-Predictive policing	-Non-predictive policing
	Context	-Place-based policing	-Person-based policing
	Outcome	-Crime reduction -Rate of crime prediction -(possibly) Efficient crime prevention(quantifiable)	-Unquantifiable literature review in terms of efficiency
Technical		Quantitative studies in : -academic journal articles -Master and doctoral dissertations -Books	-Qualitative studies -Non-academic writings in media (i.e. newspapers or journals)
		Grey Literatures : -Policy reports are written by academics or researchers or law enforcement professionals. -Governmental policy reports -Conference papers and proceedings	-Opinion pieces from the individual sphere (i.e. blogs or social media) -Policy reports or opinions not written by academics or researchers or law enforcement professionals.
		Language of study: English	Languages other than English
		Studies within 15 years (17 May 2007 – 17 May 2022)	Studies before 2007
		Obtainable	Unobtainable in full-text and abstract

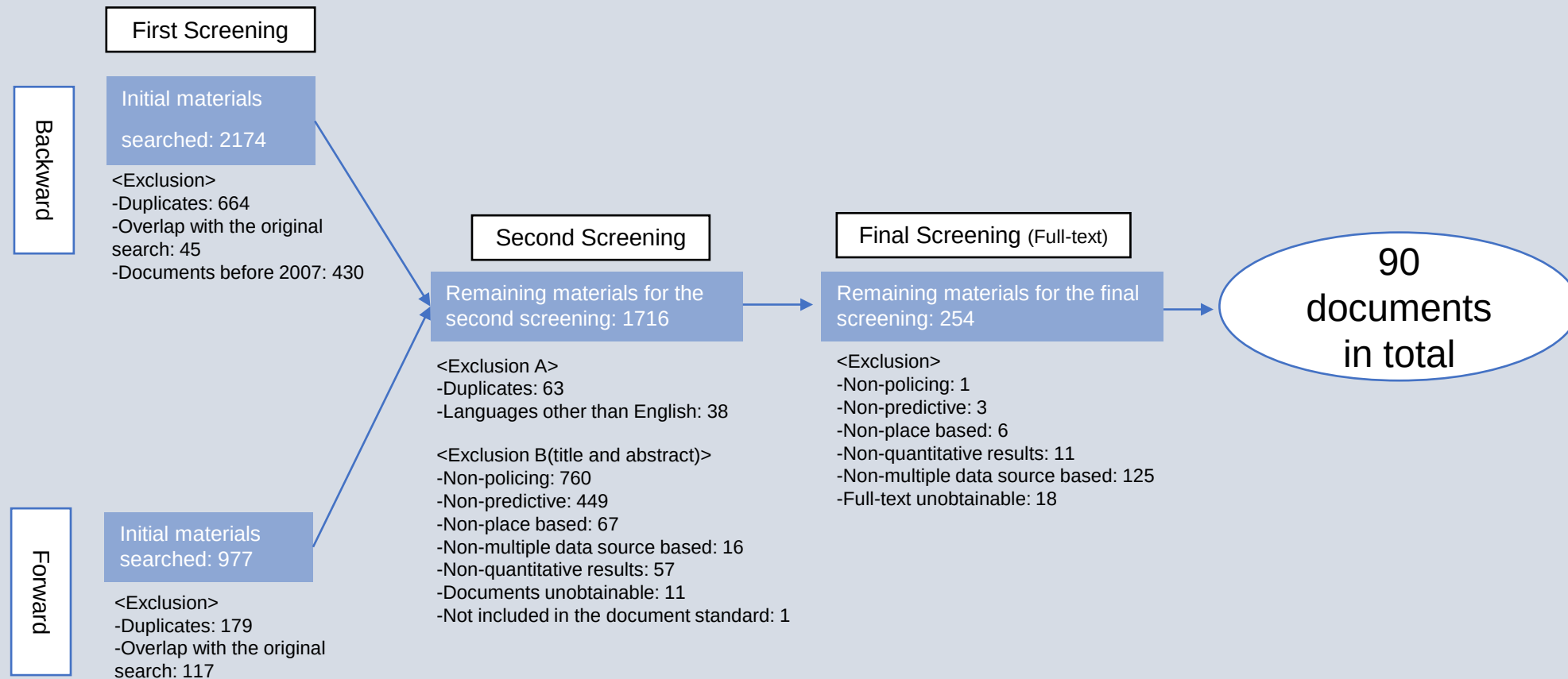
+ Update Search was conducted
(Period : 17 May 2022 ~ 7 Sep 2022)

+ Citation Chaining Search (backward and forward) was conducted

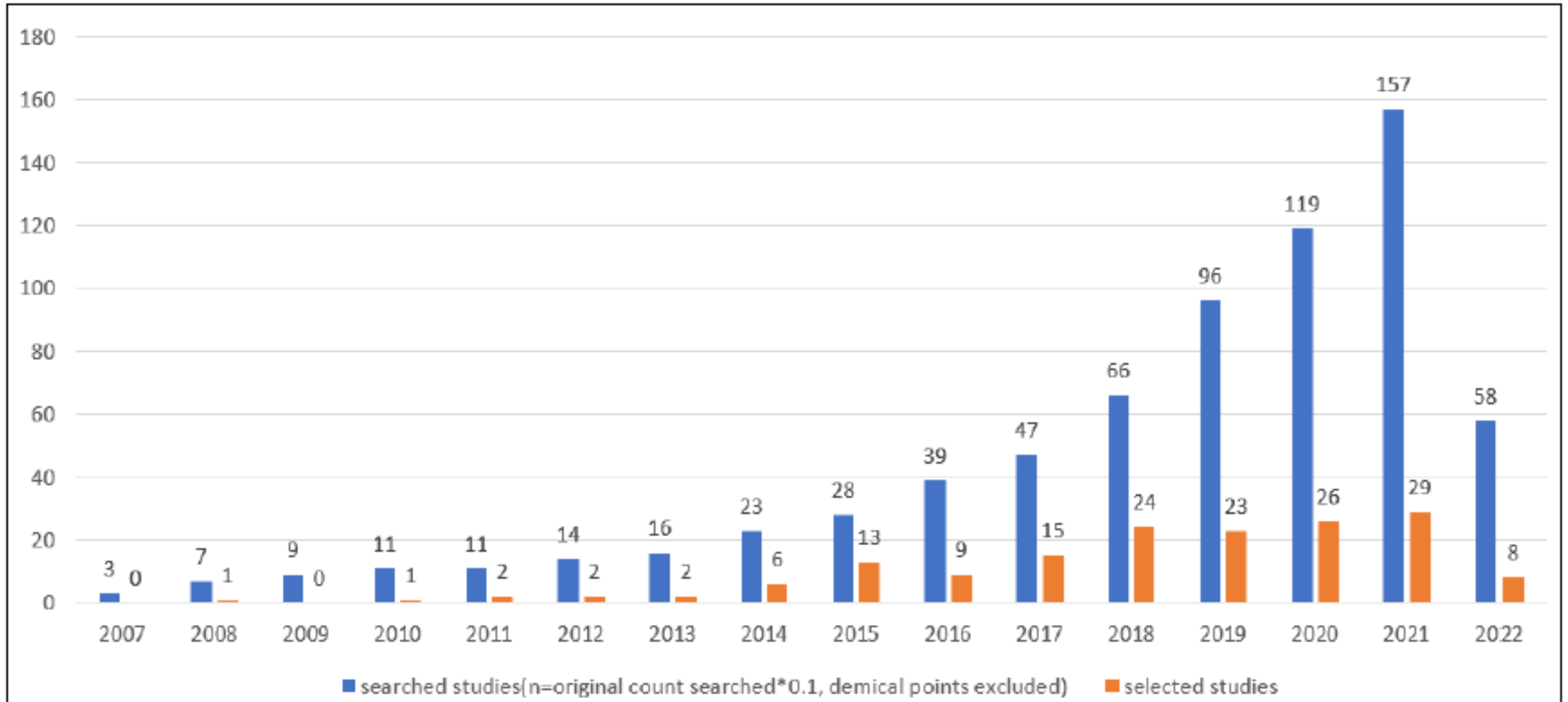
Results of basic search



Results of citation chaining search from 71 papers



Number of studies searched and selected



Categorizing the effectiveness in BDPP

Evidential Value

Type 1: Applied in the real-world testing(i.e., RCTs) with displacement effect inspection



Type 2: Applied in the real-world testing(i.e., RCTs)



Type 3: Retrospective(tested on past data) but invented novel prediction model



Type 4: Retrospective(tested on past data) with no algorithmic model construction

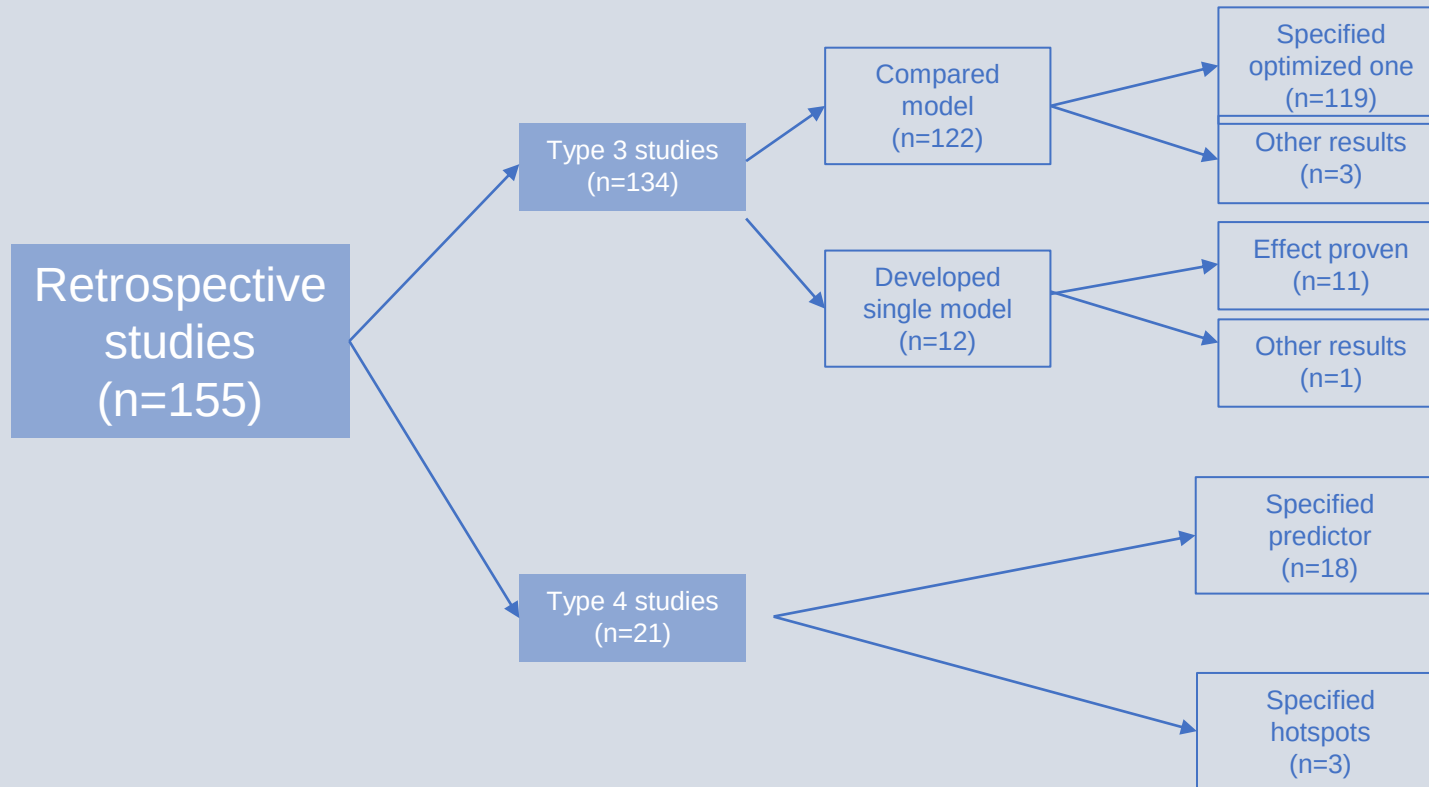


	Tested in the real-world application	Tested displacement effect	Constructed direct prediction model
Type 1	O	O	
Type 2		X	
Type 3	X		O
Type 4			X

Type 1 and 2 studies (n=6)

Type	Authors (year)	Data used for prediction	Target Crime	Police intervention	Result
1	Braga & Bond (2008)	Emergency call, various qualitative data (local place characteristics, officers' opinion etc.)	Various types of crimes	Situational interventions, Aggressive interventions (arrest, patrol, etc.), etc.	19.8 % reduction (in general, statistically significant than the control group)
1	Carter et al (2021)	Police crime data, drug overdoes data (ER data), crime cost estimation data	Social harm index	Vehicle patrol or foot patrol	The effect of the intervention was significant (p=.0295, social harm cost decreased \$38.6 per 10.4 min of policing)
1	Ratcliffe et al (2021)	Crime data, demographic data, weather data, etc.	Property and violent crimes	Officer awareness, marked car patrol, unmarked car patrol.	31% reduction only on property crime when applying marked car patrol
2	Wyatt & Alexander (2010)	Crime data (number of crimes), traffic accident data, driving under the influence data.	Traffic crashes	Vehicle stops (with visibility like blue lights)	Decreased number of fatal(15.9%) and injury(30.8%) accidents etc after the intervention.
2	Florence et al (2011)	Crime data, demographic data, health service records	Violent crimes	Targeted deployment of police resources (presence, CCTV)	Effective in hospital admissions from violence(7->5 per 100k people in treatment while 5->8 in control), and recorded wounding(54->82 per 100k people in treatment while 54->114 in control)
2	Hunt et al (2014)	Crime data, disorder calls, seasonal variations, juvenile arrests	Property crime	Directed patrol etc.	No significant effect on reducing crimes than the control group

Type 3 and 4 studies (n=155)



Type 1 and 2 studies (n=6)

Type	Authors (year)	Data used for prediction	Target Crime	Police intervention	Result
1	Braga & Bond (2008)	Emergency call, various qualitative data (local place characteristics, officers' opinion etc.)	Various types of crimes	Situational interventions, Aggressive interventions (arrest, patrol, etc.), etc.	19.8 % reduction (in general, statistically significant than the control group)
1	Carter et al (2021)	Police crime data, drug overdoes data (ER data), crime cost estimation data	Social harm index	Vehicle patrol or foot patrol	The effect of the intervention was significant (p=.0295, social harm cost decreased \$38.6 per 10.4 min of policing)
1	Ratcliffe et al (2021)	Crime data, demographic data, weather data, etc.	Property and violent crimes	Officer awareness, marked car patrol, unmarked car patrol.	31% reduction only on property crime when applying marked car patrol
2	Wyatt & Alexander (2010)	Crime data (number of crimes), traffic accident data, driving under the influence data.	Traffic crashes	Vehicle stops (with visibility like blue lights)	Decreased number of fatal(15.9%) and injury(30.8%) accidents etc after the intervention.
2	Florence et al (2011)	Crime data, demographic data, health service records	Violent crimes	Targeted deployment of police resources (presence, CCTV)	Effective in hospital admissions from violence(7->5 per 100k people in treatment while 5->8 in control), and recorded wounding(54->82 per 100k people in treatment while 54->114 in control)
2	Hunt et al (2014)	Crime data, disorder calls, seasonal variations, juvenile arrests	Property crime	Directed patrol etc.	No significant effect on reducing crimes than the control group

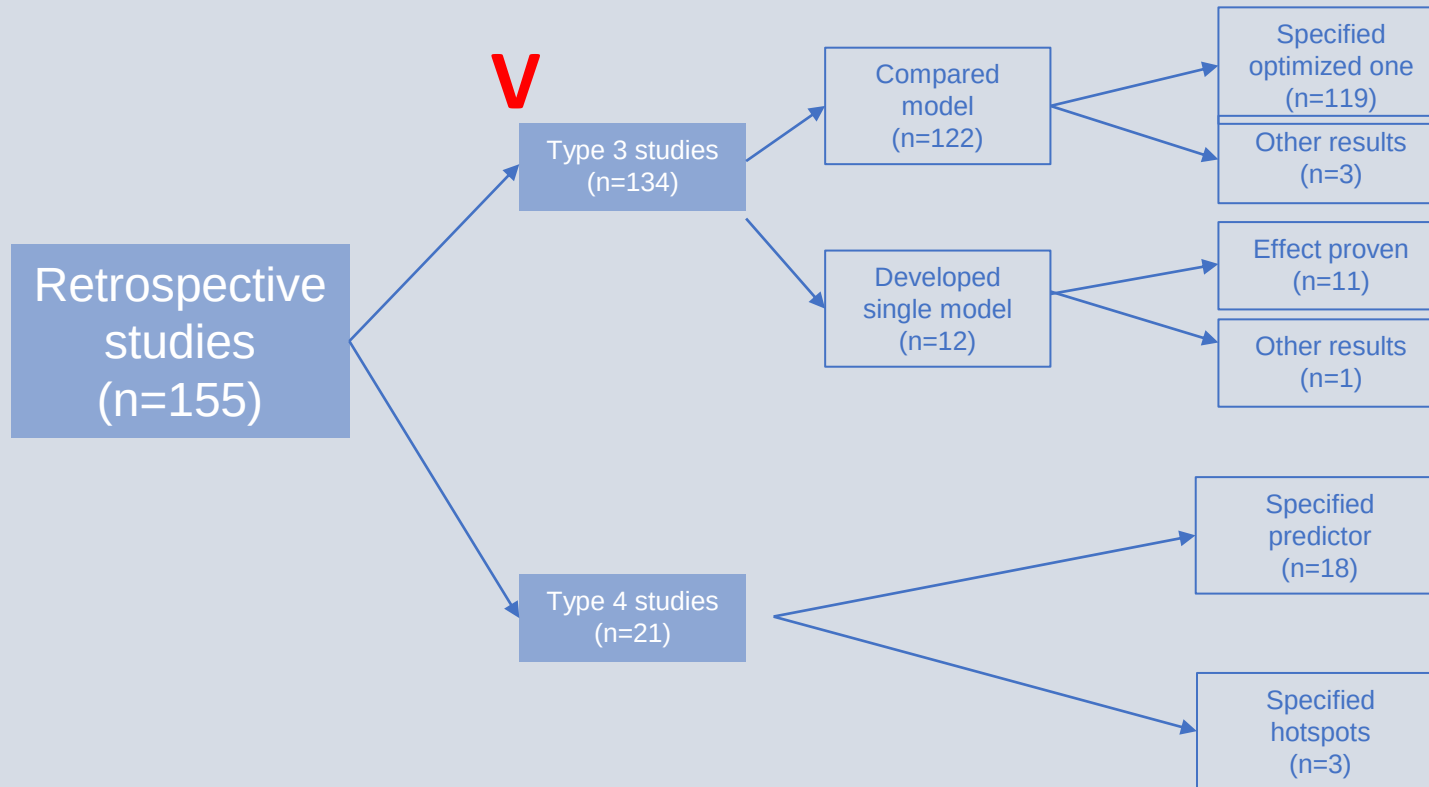
V

Type 1 and 2 studies (n=6)

Type	Authors (year)	Data used for prediction	Target Crime	Police intervention	Result
1	Braga & Bond (2008)	Emergency call, various qualitative data (local place characteristics, officers' opinion etc.)	Various types of crimes	Situational interventions, Aggressive interventions (arrest, patrol, etc.), etc.	19.8 % reduction (in general, statistically significant than the control group)
1	Carter et al (2021)	Police crime data, drug overdoes data (ER data), crime cost estimation data	Social harm index	Vehicle patrol or foot patrol	The effect of the intervention was significant ($p=.0295$, social harm cost decreased \$38.6 per 10.4 min of policing)
1	Ratcliffe et al (2021)	Crime data, demographic data, weather data, etc.	Property and violent crimes	Officer awareness, marked car patrol, unmarked car patrol.	31% reduction only on property crime when applying marked car patrol
2	Wyatt & Alexander (2010)	Crime data (number of crimes), traffic accident data, driving under the influence data.	Traffic crashes	Vehicle stops (with visibility like blue lights)	Decreased number of fatal(15.9%) and injury(30.8%) accidents etc after the intervention.
2	Florence et al (2011)	Crime data, demographic data, health service records	Violent crimes	Targeted deployment of police resources (presence, CCTV)	Effective in hospital admissions from violence(7->5 per 100k people in treatment while 5->8 in control), and recorded wounding(54->82 per 100k people in treatment while 54->114 in control)
2	Hunt et al (2014)	Crime data, disorder calls, seasonal variations, juvenile arrests	Property crime	Directed patrol etc.	No significant effect on reducing crimes than the control group

V

Type 3 and 4 studies (n=155)



Summary

The effectiveness was proven?

Compared to the high level of interest in the topic, the number of studies classified as strong research is limited, but some of them have demonstrated effectiveness.

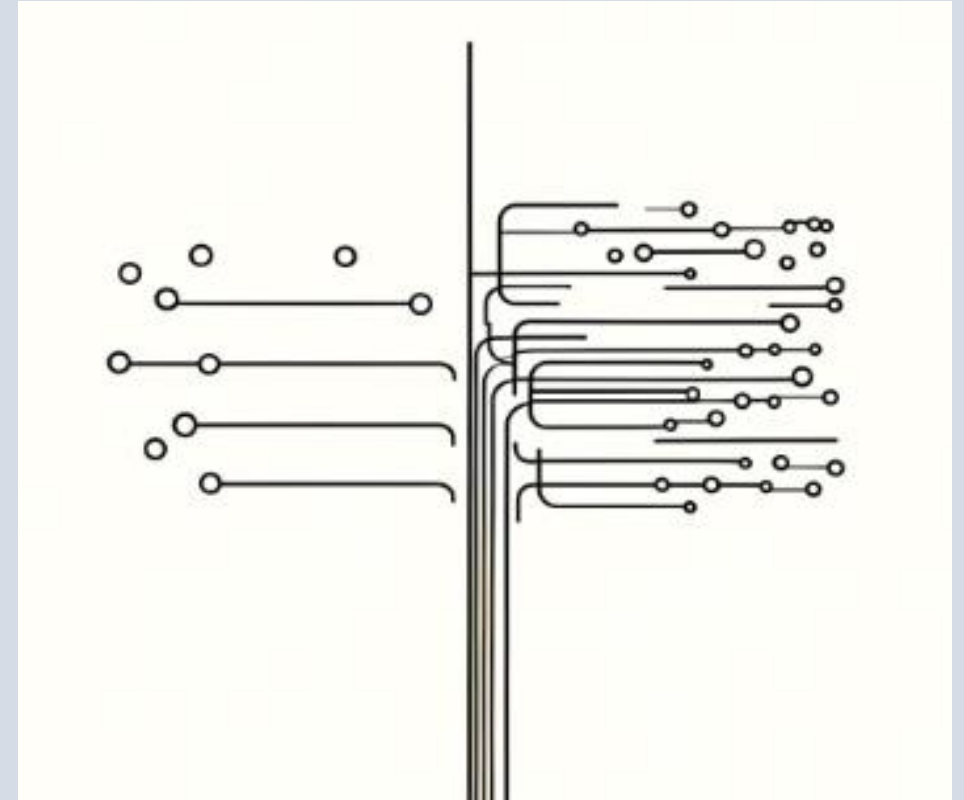
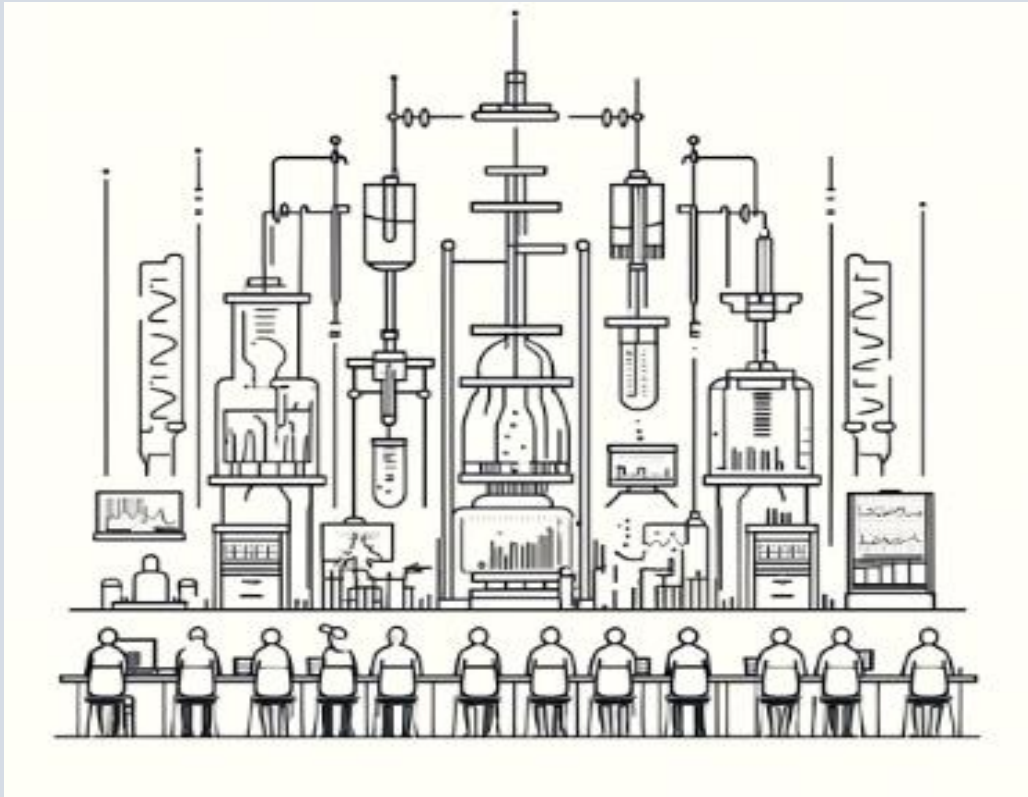
The proven effectiveness is strong enough to offset the concerns?

Not there yet. In our opinion.



Next steps required-If BDPP is to be used,

1. Further experiments are needed that meet the same experimental standards as those used in the Type 1 & 2 studies.



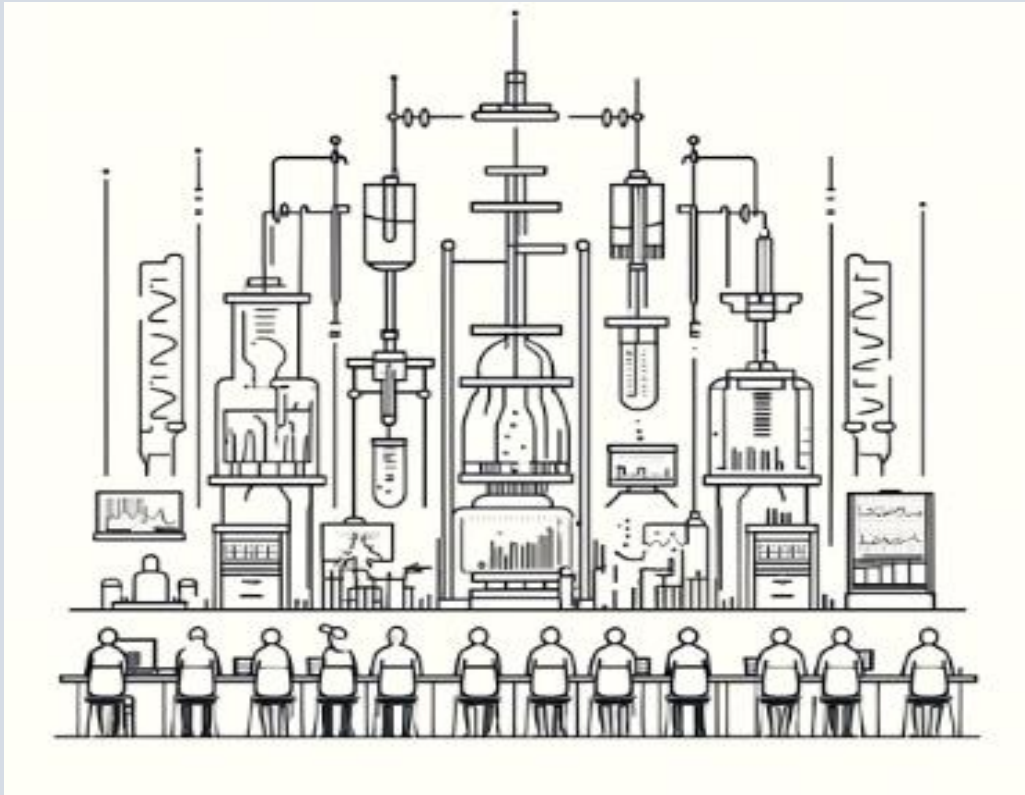
2. Comparison with an algorithmic prediction model using only criminal data is required to mitigate or avoid concerns.

Next steps required-If BDPP is to be used,

1. Further experiments are needed that meet the same experimental standards as those used in the Type 1 & 2 studies.

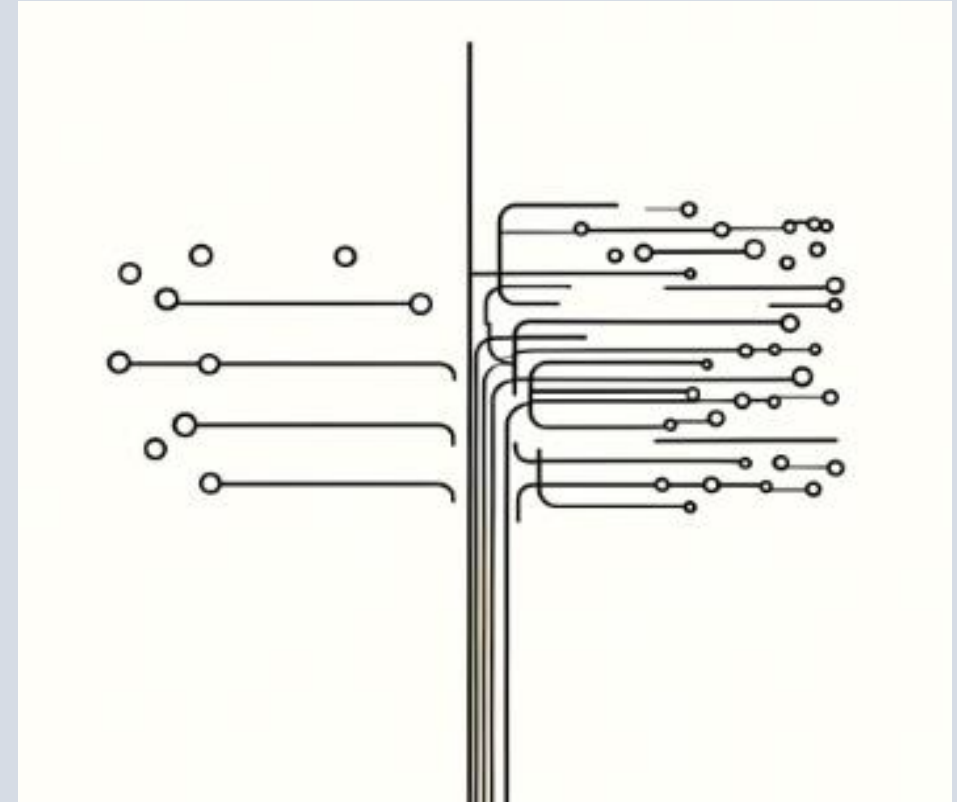
In general, RCTs(Randomised Controlled Trials) are needed. Specifically,

- Bearing in mind that predictive policing is not a silver bullet, specifically identify the target crime and policing intervention strategy
- Should test displacement or diffusion of benefits
- If the actual application at field level fails (dosage failure), there may be no significant effect. Therefore, a sensitive plan to persuade field officers would be an asset.



Next steps required-If BDPP is to be used,

- If a BDPP is to be implemented (if LEAs want access to other civil data in addition to the criminal data they already have),
- LEAs would first **need to demonstrate that the use of multi-source data leads to significant efficiencies in crime prediction over the use of criminal data alone.**
- Although the use of criminal data alone will not prevent algorithmic bias, the above logic can be applied to create a minimum threshold for police state concerns.



2. Comparison with an algorithmic prediction model using only criminal data is required to mitigate or avoid concerns.

Why we need above?: to win the community trust

-Tom Tyler's theory of Procedural Justice(PJ) was the benchmark for democratic & trustworthy policing. PJ stands for the 'good explanation with respectful attitude' from police officers to citizens

-If we apply this micro level logic of PJ to a macro level of predictive policing,

-Governments(or academics), who have more expertise and access to data than citizens, have an obligation to transparently explain the facts of predictive policing to the citizens.

-Since effectiveness(in reducing crime, and etc.) has been the basis for claiming the legitimacy of predictive policing, it should be explained to civil society at the first hand. After this review is made, weighing the benefits of predictive policing against the threats it may pose will allow for informed democratic decision-making about predictive policing(and if the social discussion is to be democratic, each party should have the same level of information about the issue under discussion)

-"What level of systematic procedural justice are we adhering to?" is the question that should be asked in the responsible democratic government to ensure the community trust



Thank you for listening!

Youngsub.lee.21@ucl.ac.uk