SYSTEMATIC REVIEW

The spatial patterning of emergency demand for police services: a scoping review

Samuel Langton[1], Stijn Ruiter[1,2*] and Linda Schoonmade[3]

Abstract

This preregistered scoping review provides an account of studies which have examined the spatial patterning of emergency reactive police demand (ERPD) as measured by calls for service data. To date, the field has generated a wealth of information about the geographic concentration of calls for service, but the information remains unsynthesised and inaccessible to researchers and practitioners. We code our literature sample (N = 79) according to the types of demand studied, the spatial scales used, the theories adopted, the methods deployed and the findings reported. We find that most studies focus on crime-related call types using meso-level (e.g., neighborhood) spatial scales. Descriptive methods demonstrate the non-random distribution of calls, irrespective of their type, while correlational findings are mixed, providing minimal support for theories such as social disorganization theory. We conclude with suggestions for future research, focusing on how the field can better exploit open data sources to 'scale-up' analyses.

Keywords Police, Calls for service, Demand, Geography, Spatial

Introduction

Public-initiated demand for police services is referred to as 'reactive' demand. Reactive demand can be generated in a number of ways, such as an in-person request for assistance out on the street, the reporting of a crime (e.g., at a police station), or as an emergency call for service (e.g., 911). The police then, if it is deemed needed, *respond* to the demand by supplying their services. Since the advent of computerized dispatch records, police demand has been most commonly measured using emergency calls for service data (Laufs et al., 2021). The characteristics of these data are unique. Typically, the time that the call was received and the call and/or incident location are automatically logged in the computerized dispatch system. Calls are logged irrespective of their seriousness, and irrespective of their criminal nature. The existence of such data, and their accessibility through open data licences, has sparked a wealth of research into the spatial (and in some cases, temporal) characteristics of public demand for police services, including non-crime forms of public demand (e.g., mental health crises). This preregistered (https://osf.io/5zshd/) scoping review aims to provide a descriptive account of studies which have sought to describe and/or explain the spatial patterning of emergency reactive police demand ('ERPD') as typically measured using emergency calls for service data.

The motivations for conducting this review are threefold. First, a wealth of theoretically-grounded (and atheoretical) research has been published in an effort to identify the major (spatial) correlates of emergency calls for service. In doing so, researchers have built a considerable body of evidence within which we might identify 'empirical regularities'. Yet, no attempt has been made to synthesise these findings. Second,

*Correspondence:
Stijn Ruiter
s.ruiter@uu.nl
[1] Netherlands Institute for the Study of Crime and Law Enforcement, Amsterdam, The Netherlands
[2] Department of Sociology, Utrecht University, Utrecht, The Netherlands
[3] Medical Library, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

recent research has demonstrated that the public rely on the police for both crime-related and non-crime-related services (Langton et al., 2022; Ratcliffe, 2021) but it remains unclear to what extent these different (non-criminal) demand types are featured in spatially-sensitive research. Third, place-based crime researchers advocate for the usage of fine-grained spatial scales (Steenbeek & Weisburd, 2016; Weisburd, 2015; Weisburd et al., 2008) but the extent to which the police demand literature reflects this development is unclear. Similarly, the usage of different temporal scales (if any) remains unknown. In addition to these primary motivations, the scoping review covers a wide array of other attributes in the literature that could be of use to academics and practitioners, including the methods deployed (descriptive and explanatory), the justifications for each study, and the study regions and study periods covered.

With these motivations in mind, the following research questions are posed:

- RQ1. How have authors justified the study of ERPD?
- RQ2. What types of ERPD have been studied?
- RQ3. At what temporal and spatial scale has ERPD been studied?
- RQ4a. What methods have been used to describe the spatial and temporal patterning of ERPD?
- RQ4b. What descriptive findings have been reported?
- RQ5a. What theoretical frameworks have been used to explain the patterns observed?
- RQ5b. What methods have been used to explain the patterns observed?
- RQ5c. What explanatory findings have been reported?

The scoping review is structured as follows. First, we provide a comprehensive overview of the literature collection process. This includes the search criteria, database searches, eligibility criteria and screening. Here, we highlight any deviations or additions from the preregistration protocol. We then provide a breakdown of the literature collated from this process. Second, we detail the methods used to code the literature according to themes which correspond to the research questions posed or additional attributes of interest (e.g., study regions). Third, we outline our findings based on the synthesis of the coded literature. We conclude with a broad overview of our descriptive account and provide suggestions for future research. The data and code used for the scoping review have been made openly available for scrutiny, reproduction and reuse (https://osf.io/5zshd/).

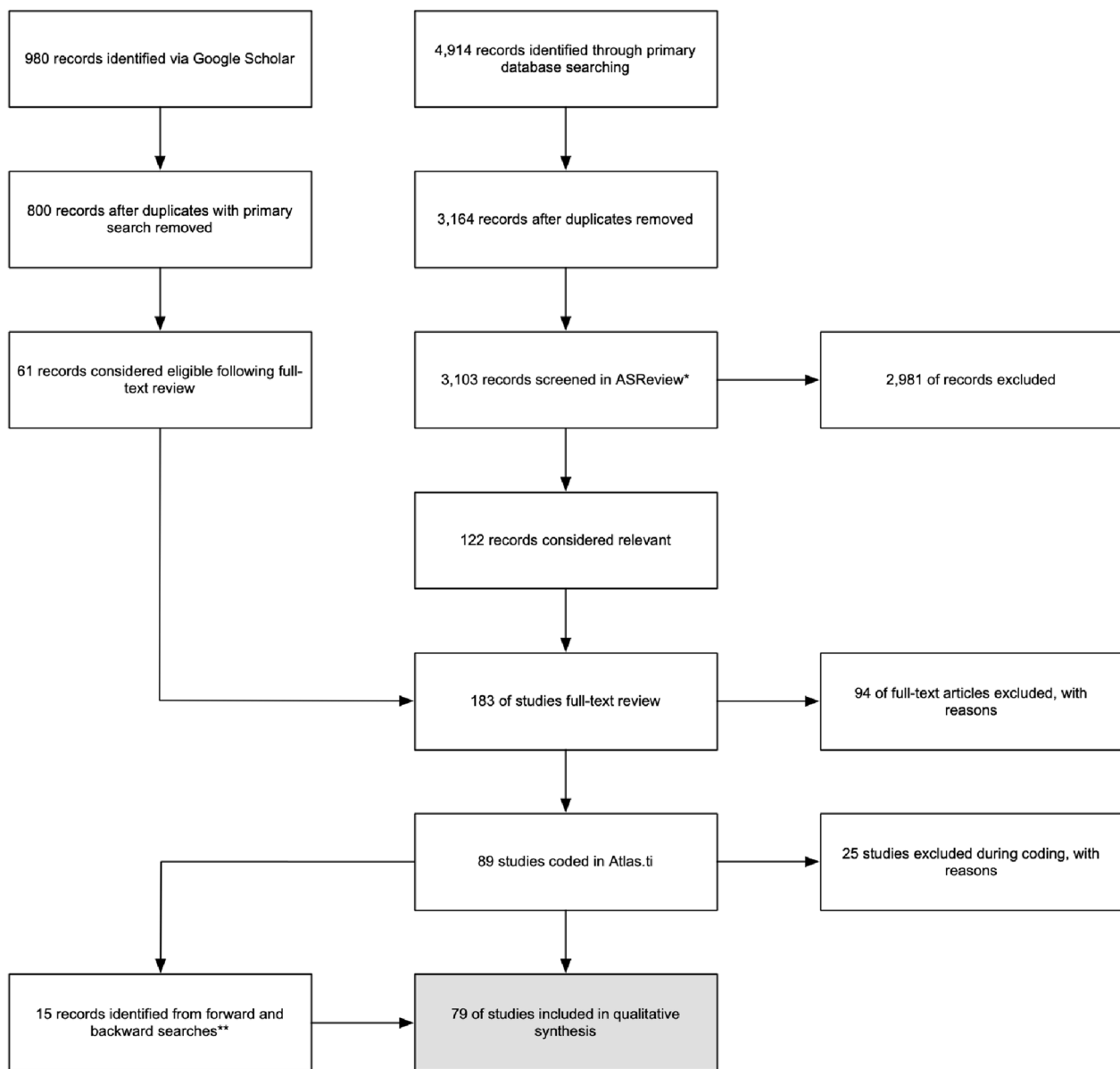## Data and methods
Broad scope
We focus our attention on studies that have examined the spatial patterning of emergency calls for service as a measure for ERPD. We know from a cursory overview of existing research that a small number of locations tend to be responsible for a disproportionately large amount of calls for service that in turn consume a disproportionately large amount of police deployment time (Langton et al., 2022; Ratcliffe, 2021). We also know that these locations can have particular correlates which are thought to be responsible for generating the high demand (Boulton et al., 2017). Consequently, studies that have sought to describe and/or explain this spatial patterning are thought to hold considerable value.

As such, the spatial component in this review is essential, defined as those studies with a micro or meso-level geographic unit of analysis such as a street segment or neighborhood (Weisburd, 2015). Studies that examine city or nationwide trends on aggregate, for instance, are not included. The temporal component is not essential: cross-sectional *and* longitudinal studies with a micro or meso geographic unit of analysis are included. In this way, we hope to provide a comprehensive evaluation about the extent to which spatial examinations of ERPD are temporally-sensitive. We consider any output (e.g., book, book chapter, journal article, report) published in English.

Three key research areas which often fall within the above definition are not included in the scoping review. First, we exclude intervention studies. For instance, experimental studies which evaluate the effectiveness of hotspot policing interventions. We deem this subfield to be distinct and it has recently been the subject of systematic review (Braga et al., 2019). Second, we exclude those studies which focus on simulation or queuing models. Third, we exclude those studies which examine extreme events, such as a natural disaster or COVID-19. In doing so, we focus on *descriptive* and/or *correlational* studies. We return to this decision (and the challenges it comes with) in the discussion.

Primary database searches
A comprehensive search was performed in the bibliographic databases Scopus (via Elsevier), Web of Science Core Collection (via Clarivate), Criminal Justice Abstracts (CJA, via EBSCO) and the International Bibliography of Social Sciences (IBSS, via ProQuest) from inception to August 25, 2021. The following terms were used (including synonyms and closely related words) as index terms or free-text words: 'police' and 'call for service'. Broader terms which encompass calls for service,

Langton *et al. Crime Science* (2024) 13:1

Page 3 of 16

980 records identified via Google Scholar

4,914 records identified through primary database searching

800 records after duplicates with primary search removed

3,164 records after duplicates removed

61 records considered eligible following full-text review

3,103 records screened in ASReview*

2,981 of records excluded

122 records considered relevant

183 of studies full-text review

94 of full-text articles excluded, with reasons

89 studies coded in Atlas.ti

25 studies excluded during coding, with reasons

15 records identified from forward and backward searches**

79 of studies included in qualitative synthesis

* 43 studies without abstracts screened manually.
** 16 studies identified but one study excluded during coding, with reasons.

Fig. 1 Flow diagram depicting the literature search results numbers resulting in the final yield for coding

such as 'public demand' and 'reactive demand', were also included. The search was performed without date or language restrictions, although we subsequently only included those published in English. The specific search terms used and results for all primary databases were reported in the preregistration document and can be found in the Appendix. The design of the search terms,

searches themselves and subsequent cleaning (e.g., de-duplication) were carried out by a librarian (author 3).

Screening
As detailed in Fig. 1, the primary database searches yielded 3,164 results after duplicates were removed. Screening of these articles was conducted using

Langton *et al. Crime Science*   (2024) 13:1

Page 4 of 16

Table 1 Inter-rater reliability descriptive statistics for abstracts screened in ASReview

	Rater 1	Rater 2
n uploaded	3103	3103
n reviewed	489	535
% reviewed	15.76	17.24
n flagged relevant	115	82
n flagged irrelevant	374	453
% flagged relevant	23.52	15.33
n unreviewed	2614	2568
n total irrelevant	2988	3021
% total relevant	3.71	2.64
n relevant v. irrelevant	39	5
n relevant v. unreviewed	0	1

ASReview (see Schoot et al., 2021). Using the title and abstract, ASReview uses active learning methods to present users with the most relevant literature in an ordered sequence. The algorithm is initially informed by a small selection of 'relevant' and 'irrelevant' articles selected by the user, and is then constantly updated with the subsequent screening decisions. Given the requirement for a title and abstract, 43 documents with incomplete information were removed for manual review. This left 3,103 unique documents for screening in ASReview. Judgements on whether literature were 'relevant' or 'irrelevant' were made on criteria stated in the pre-registration document. Here, we added an additional criteria based on a pilot round of ASReview screening: as noted, we exclude studies focusing on extreme events (e.g., COVID-19) (Table 1).

Both first and second authors conducted screening in ASReview. While both raters began with the same RIS file containing the 3,103 references, the initial selection of articles for kick-starting the algorithm was independently selected by each rater. Each rater stopped reviewing when one hundred articles in a row were screened as 'irrelevant'. Descriptive statistics on the number of abstracts reviewed, the decisions made, and the extent

Table 2 Summary of main effect directions by broad classification of predictor variables on call volume

Predictor variable class	Negative %	None %	Positive %	Times tested
Collective efficacy				
Collective efficacy	64	36	0	11
Community institution	11	89	0	18
High risk population				
Single persons	0	0	100	9
Males	7	79	14	14
Young people	28	45	28	29
Opportunity/routine activities				
Alcohol outlets	8	46	46	24
Ambient population (density)	17	67	17	24
Points of interest	10	40	50	50
Other/mixed				
Foreclosures	11	22	67	9
Vacant housing or demolitions	12	65	24	17
Crime or calls for service (various)	2	13	84	89
Social disorganization				
Resident immigrants	0	75	25	8
Physical disorder	0	33	67	9
Unemployment	0	23	77	22
Family disruption	15	58	27	26
Ethnic diversity	25	66	9	32
Resident ethnicity	12	61	27	33
Education level	30	45	25	44
Population turnover	22	45	33	64
Resident population (density)	35	25	40	65
Deprivation	21	48	31	121

of agreement, are summarized in Table 2.[1] This demonstrates how, despite both raters beginning with the full reference list (N=3103), the raters only reviewed 16% (N=489) and 17% (N=535) of the total literature sample before hitting one hundred irrelevant articles in a row. By this point, only one article had been flagged as 'relevant' that remained unreviewed (i.e., beyond the one hundredth irrelevant article) by the other (in Table 2: *n. relevant v. unreviewed*). This article was later excluded as irrelevant. These steps strongly indicated that the literature sample had been thoroughly scanned for relevant articles, and that no more relevant articles remained.

That said, there were some disagreements before that end point. For example, Rater 1 flagged 39 studies as relevant that Rater 2 flagged as irrelevant, and conversely, Rater 2 flagged five studies as relevant that Rater 2 flagged as irrelevant. Nevertheless, an inter-rater reliability test confirmed a reasonable level of agreement (Kappa=0.715, p<0.001). Besides, studies identified as 'relevant' by *either* first or second authors were used for full-text review by the first author. A complete report on the inter-rater reliability in ASReview is available online (https://osf.io/5zshd/).

Secondary searches
Google scholar
The primary searches were supplemented with two secondary sources, namely, Google Scholar and forward/backward searches. Google Scholar serves as a check following the primary searches and as a useful search engine for obtaining grey literature (Haddaway et al., 2015). To permit bulk downloading of the search, we made the Google Scholar query using the *Public or Perish* software (Harzing, 2010). Searches were limited to 1000 results. A simplified version of our advance search terms (see Appendix) was carried out on 29 September, 2021. The literature obtained from this search (N=61) were subject to a full-text scan to determine their relevance.

Forward/backward searches
Forward and backward searches were conducted by two research assistants after the final list of 'relevant' articles had been established based on the reference lists of these relevant articles. The first author reviewed those flagged as relevant and made the final decision on their inclusion.

Literature yield
A complete summary of the literature yield from our searches is summarized in Fig. 1. This includes the various de-duplication and exclusion phases, culminating in the 79 articles processed for full review and coding.

Coding
Relevant literature was coded according to themes in Atlas.ti (version 8). Code themes were created either in direct correspondence to a research question or as part of the data charting, which included information such as the study region and study period. The main codes were used as detailed in the preregistration document, although some additional codes were added to reflect the data charting (e.g., study regions), to capture the in-text description that stated (or indicated) that 'emergency calls for service' data were being used, or to assist in the search itself (e.g., a code for excluded literature). The entire Atlas.ti project bundle has been archived and is available upon request.[2] The coding scheme and quotations for each article as coded in Atlas.ti were exported as spreadsheets, and along with the corresponding R scripts (R Core Team, 2022), are openly available (https://osf.io/5zshd/).
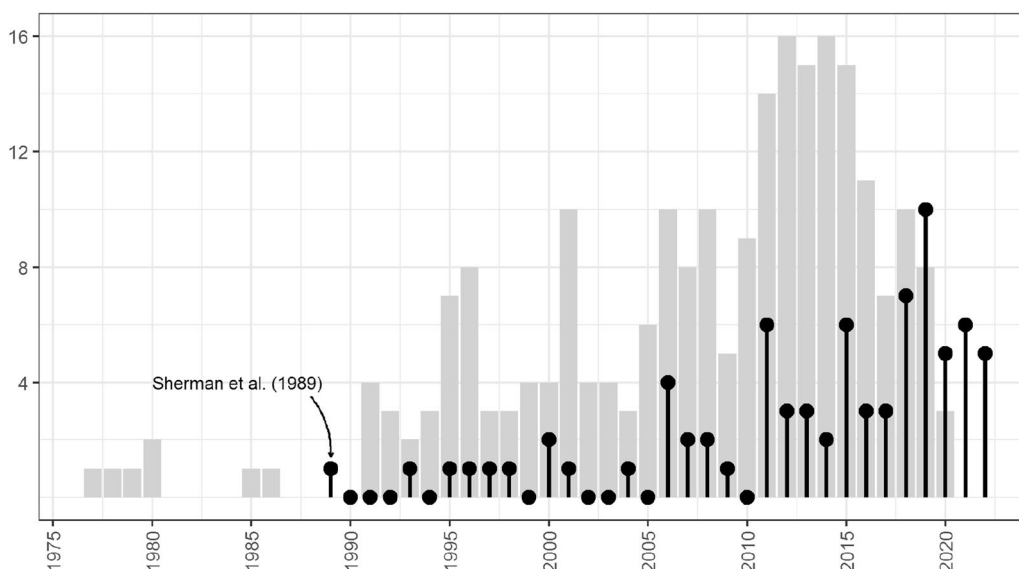
Findings
Basic information
The publication frequency and corresponding data coverage over time are visualised in Fig. 2. Interest in the spatial patterning of ERPD, as measured through calls for service data, can be traced back to Sherman et al. (1989). This followed the adoption of Computer Aided Dispatch (CAD) systems and the corresponding archiving of data in the United States (Pierce et al., 1988). A number of studies followed suit (Spelman, 1995; Warner & Pierce, 1993) and in recent years the field has clearly been on an 'upward' trajectory in popularity, even when excluding studies examining the impact of the COVID-19 pandemic. Consequently, data coverage has concentrated around 2005–2015 during which time some countries were experiencing the tail-end of macro-level declines in crime, but increases in other forms of police demand (CoP, 2015).

Our literature sample overwhelmingly features study regions in the United States (Fig. 3). Overall, largely due to the focus on the US and Canada, there is a clear skew towards English-speaking countries. This might, in part, simply be because we only include those studies published in English. It might also be attributable to data access, particularly due to the widespread availability of open data portals from US police departments. Despite this, we note that open data access in the US has

---

[1] These descriptives were generated using an automated reporting tool, created as part of this study, which is openly available (https://anonymous.4open.science/r/asreview_irr-045C/README.md).

[2] The project bundle cannot be uploaded to a public repository because it contains all the relevant PDF files for each article, many of which are subject to publishing copyright laws.

Langton *et al. Crime Science* (2024) 13:1

Page 6 of 16

Fig. 2 Publications year frequencies (in black) and the corresponding study period coverage (in grey) for the literature sample
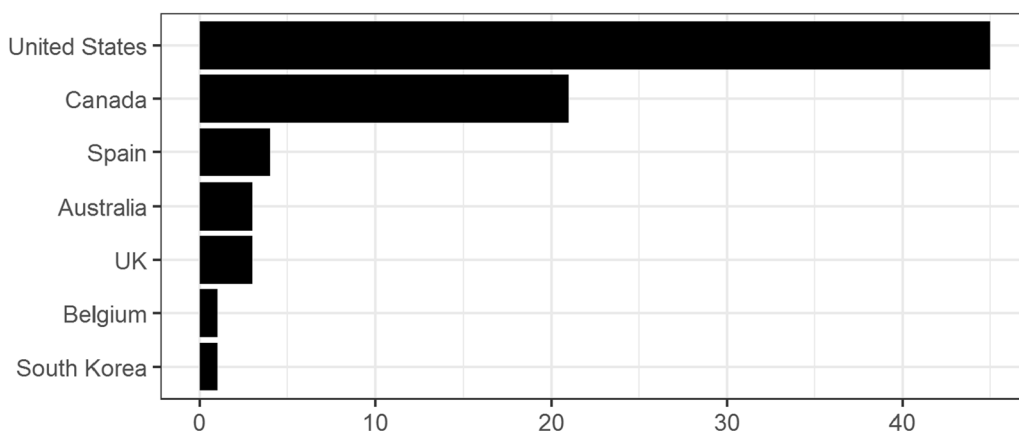

Fig. 3 Study region countries

not translated into open science: no study in our literature sample provided enough materials (data under open licence, code in open software) to reproduce findings in their entirety.

Study justifications

The justifications that researchers have used to study the spatial patterning of ERPD are diverse. A key reason is *theory testing*. This has typically come in four forms, namely, (1) testing old theories with new data—e.g., social disorganization but using calls for service as a new measure for crime (Warner & Pierce, 1993), (2) testing old theories with new methods—e.g., social disorganization using advanced spatio-temporal techniques (Lymperopoulou et al., 2022), (3) testing combinations

of theoretical frameworks—e.g., routine activities with social disorganization (Andresen, 2006a), and (3) testing the appropriateness of existing theories in a different context, such as suburban areas (Roh & Choo, 2008), non-Western countries (Kim & Kim, 2022) or under-researched call types such as domestic violence (Roman & Reid, 2012) or mental health (Vaughan et al., 2016).

Researchers have also justified their endeavours after having observed an *emerging demand problem* such as a recent increase in a specific call type, for instance, burglary during the 1990s (Guidi et al., 1997) or more recently, mental health crises (Hodgkinson & Andresen, 2019; Koziarski, 2021). There can also be an *emerging demand-generating problem*, such as foreclosures following a housing crisis (Pfeiffer & Lucio, 2015) or Airbnbs
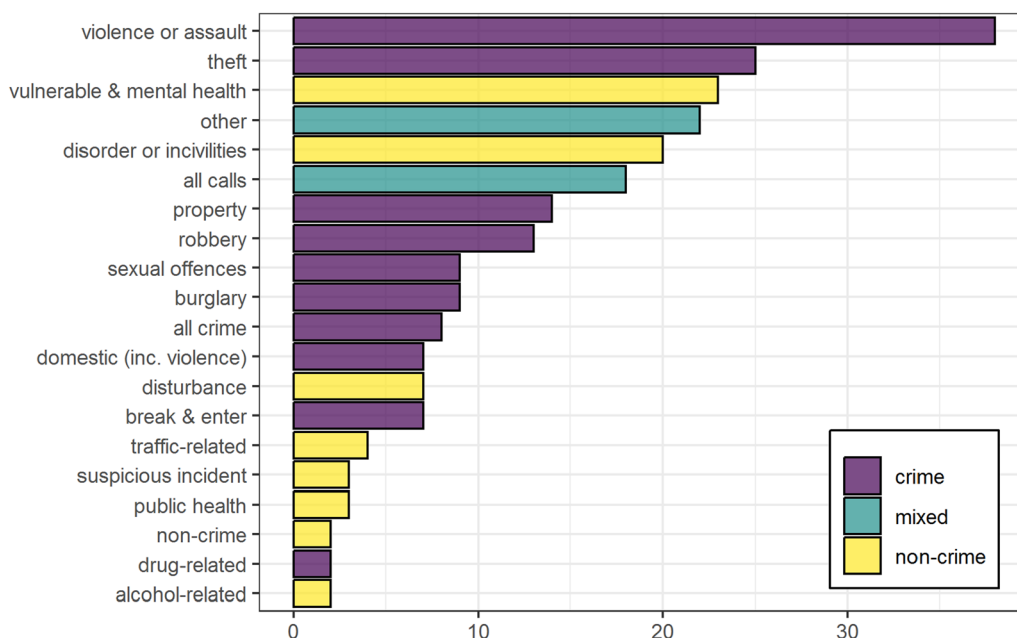
Fig. 4 Call types frequency following simplification. Note that some studies examined multiple call types

(Holm & Monaghan, 2021) for which researchers are motivated to test its co-existence with calls for service volume.

In our literature sample, justifications were also data or measurement orientated. This included assessing the *impact of spatial scale* when testing longitudinal stability (Andresen & Malleson, 2011) and the consequences of aggregating across crime types (Andresen & Linning, 2012). The typical reliance on cross-sectional data when studying the spatial patterning of calls for service means that studies justify their research by incorporating a *longitudinal* component in the outcome variable—e.g., seasons (Marco et al., 2017) or hours of the day (Luan et al., 2016). A small number of studies have been motivated to scrutinise the use of raw call counts or resident population denominators in the outcome variable, instead focusing on the impact of using the *ambient population as the denominator* when investigating spatial concentration (Andresen, 2011b) or theory testing (Andresen, 2006b).

Demand types
As visualized in Fig. 4, the majority of call types used in analyses were related to crime (56%). Just 27% were non-crime incidents, such as those involving vulnerable people/mental health, and 17% were mixed aggregations. The most common call type by a considerable margin was violence or assault. We would highlight that drug-related

call types were only included if the author(s) specifically stated that the calls were public-initiated. Almost every study used count aggregations of calls to their spatial unit of analysis to reflect 'volume'. There were some exceptions. Some studies adjusted the counts according to a denominator, such as the ambient and/or resident population (Andresen & Brantingham, 2007; Andresen, 2006a), used counts to calculate a Location Quotient (Koziarski, 2022; Vaughan et al., 2018), or used clusters generated from a Local Moran's I analysis (Andresen, 2011a). One study calculated dispatched deployment time as a measure of supply to meet call demand, in addition to a count measure (Ellison et al., 2021). We include all studies collectively in most of the subsequent results, with the exception of the correlational findings, detailed below.

Scales
Spatial
The literature sample has been limited to those examining meso and micro level spatial scales. These are summarized in Fig. 5a. 61% of all the spatial units of analysis used were categorised as meso-level and 39% as micro-level. The operationalisation of these scales varied considerably. Generally speaking, meso-level units were defined according to census units, which given the North American-centric sample, consist primarily of blocks, tracts or dissemination areas (e.g., Koziarski, 2022; Louis
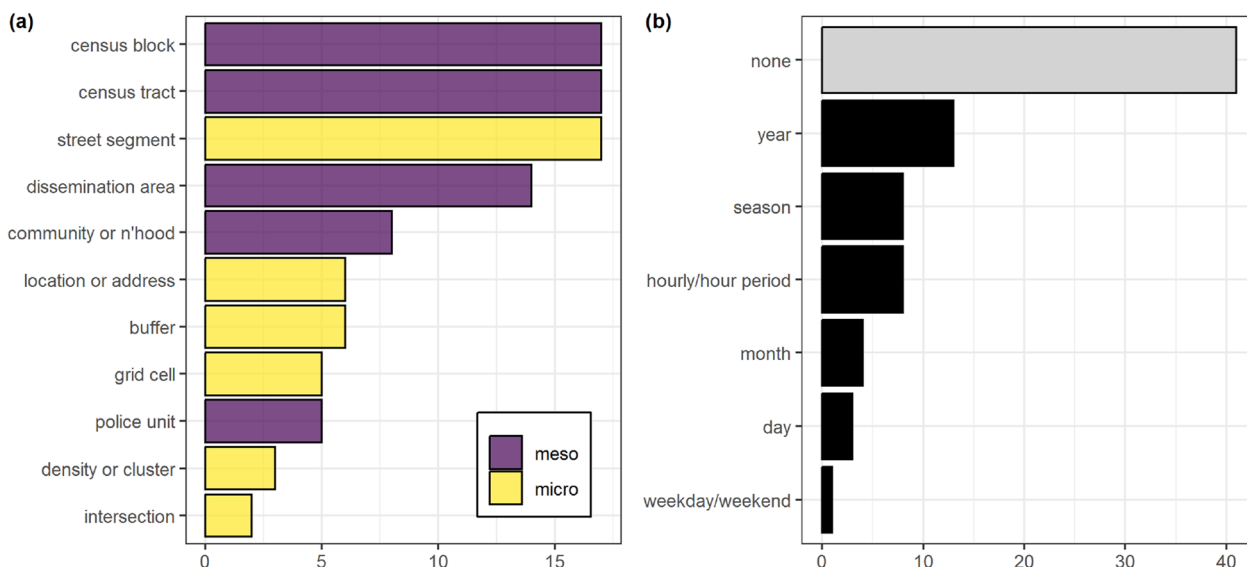
Fig. 5 Granularity of (**a**) spatial and (**b**) temporal scales. Note that some studies used more than one scale

& Greene, 2020; Reinhard, 2023).[3] Micro-level analyses were mostly conducted using street segments (Hodgkinson et al., 2020; Vaughan et al., 2016), although a number of studies have used specific addresses (e.g., Sherman et al., 1989) or synthetic grids (e.g., Clare et al., 2019). Five studies used a meso-level unit of analysis which is designed and/or used by the police jurisdiction itself (e.g., Chohlas-Wood et al., 2015; Quick, 2019). Around one-third of papers in our sample used multiple spatial scales in the same study, usually for the purposes of a sensitivity check (e.g., Andresen & Linning, 2012).

Temporal
Approximately half of studies in our sample used a cross-sectional measure of ERPD (see Fig. 5b). Of those that had a temporal component, the year was the most common temporal scale. This inevitably narrows the scope of studies to *long-term* change, such as describing longitudinal stability in demand concentration (Andresen & Malleson, 2011). It differs from studies which examine *cyclical* change, such as *seasonal* fluctuations using monthly or seasonal aggregations (e.g., Pfeiffer & Lucio, 2015), or fine-grained temporal scales such as *hours or times of the day* (Ellison et al., 2021). A small number of studies examined partitions of the week, such as days or weekdays/weekends (e.g., Bocker Parks, 2015).

---
[3] Here, we note that some categorization decisions were arbitrary for the purposes of the summary. For instance, Lower Super Output Areas are commonly referred to as 'neighborhoods' in the literature, and were categorised as such, but are census geographies in the United Kingdom.

Descriptives
Methods
We can summarize the descriptive methods used according to two principal motivations, namely, to describe the extent of *spatial concentration* or *spatial autocorrelation* in calls for police service.

The spatial concentration of calls for service has been typically described using visualizations such as kernel density maps (e.g., Andresen & Brantingham, 2007; Jones et al., 2019) or choropleth maps of raw counts aggregated to the spatial unit of analysis (e.g., Quick, 2019; Reinhard, 2023). These methods tend to only provide context to the study region based on an eye-ball assessment of the concentration. For instance: "Map 2 (violent crimes) shows that the distribution of crime in Vancouver is far from random or uniform" (Andresen & Brantingham, 2007, p. 7). Instead, or in addition, a number of studies quantify this concentration by summarising the cumulative percentage of incidents versus the cumulative percentage of spatial units (e.g., street segments). This can be visualised using a Lorenz curve (e.g., Clare et al., 2019) or by an arbitrary threshold along the Lorenz curve, such as the percentage of street segments accounting for 50% of crime (Andresen & Malleson, 2011). The Gini coefficient has been commonly used as a global statistic of concentration, summarising area under the Lorenz curve in a single number (e.g., Koziarski, 2021). The arbitrary cumulative thresholds and Gini coefficients have the advantage of comparability between study regions, which in turn help to form a consistent evidence-base and the discovery and empirical regularities, as demonstrated by what Weisburd (2015) coined the 'law of crime concentration'.

Langton *et al. Crime Science*   (2024) 13:1

Page 9 of 16

The usage of typical descriptive statistics (e.g., mean, standard deviation, range) of calls across the spatial units of analysis in a study region is common. While these statistics demonstrate the skew in the distribution (in effect, concentration) the findings are never written about in such terms and instead form the basis for pre-analysis descriptives that are rarely discussed in any detail (e.g., Boggess & Maskaly, 2014; Louis & Greene, 2020). One circumstance in which a measure of central tendency (e.g., the mean) is used and discussed in detail is when summarising the results from trajectory analysis. In this scenario, the mean number of calls are plotted to summarize each cluster's trajectory over time with no information about the variation around that mean (Hibdon et al., 2017).

Methods of determining the extent of spatial autocorrelation have either been used as a way of determining overall, citywide clustering (or dispersion) using a *Global Moran's I* statistic (e.g., Andresen, 2011b), or relatedly, as a means of identifying localized clustering of (dis)similar values using the *Local Moran's I* (e.g., Dewinter et al., 2022) or *Getis-Ord Gi** (e.g., Bocker Parks, 2015) which can then be plotted visually on a map. The former provides a straightforward test to reject the null hypothesis of randomness: a precursor to more complex analysis which seeks to explain the observed patterns (e.g., Ellison et al., 2021). The latter is presented as a substantive finding in its own right, sometimes framed as 'hotspot analysis' (Lersch & Christy, 2020) or outlier identification (Dewinter et al., 2022), to identify clusters of high-demand areas.

Findings
Without exception, descriptive statistics have demonstrated that emergency demand for police services concentrate in space, irrespective of the spatial unit of analysis, the study year, the study region or the call type under examination. That said, the extent of this concentration can vary between study regions and demand types. Generally speaking, the degree of concentration at the street segment level, quantified as the percentage of street segments which account for 50% of calls of any given type, can vary between < 1% and 8%. For those using yearly longitudinal data, the degree of concentration tends to be remain fairly stable over time (Andresen & Malleson, 2011). The bandwidths are certainly comparable to the evidence underpinning the law of crime concentration (Weisburd, 2015), indicating that a similar phenomenon exists in crime-related emergency demand. The field is currently lacking sufficient longitudinal studies of concentration using non-crime police demand (e.g., mental health) to make such a claim. The degree of global spatial autocorrelation is consistently positive,

indicating that similarly high and/or low demand areas are geographically proximal to one another. That said, visualizations of local cluster measures do indicate that dissimilar areas can be geographically proximal, supporting the push for fine-grained units of analysis that unmask such variation, rather than aggregating information away.

Explanations
Theory
Over half of studies draw upon a theoretical framework (58%), even if analyses were descriptive. The most common theory deployed, either for the explicit purpose of testing hypotheses or as a discussion point to findings, was social disorganization theory (see Fig. 6). A number of the frameworks or theories used fall within the crime opportunity theory, such as the routine activity approach, situational crime prevention, rational choice theory, and those speaking broadly about what we have termed 'risky facilities', which includes theoretical discussions on attractors and generators of crime (e.g., Lersch & Christy, 2020). A number of studies simply referred broadly to opportunities or opportunity theory (Roman & Reid, 2012). While there is some commonality in the theories used, there is considerable diversity in the frameworks deployed. The 'other' category, defined arbitrarily as those theories used three times or less in the literature sample, included dual-process theory (Hagan et al., 2018) and Klinger's Ecological theory (Taniguchi & Salvatore, 2018), among others. 42% (N = 36) of studies did not explicitly state a theoretical framework and were classified as 'atheoretical'. This subset includes studies which deployed explanatory models which most readers would recognise as derived from a particular theory, but without explicit reference to it (Ellison et al., 2021; Marco et al., 2018) but also descriptive contributions with no theory-testing goals (Ratcliffe, 2021).

Methods and models
The field appears to have progressively moved away from rudimentary statistical techniques (e.g., OLS regression, bivariate correlations) as the primary means of explaining the spatial patterning of calls for service. With the consistent finding of spatial autocorrelation, and the increased availability (and usability) of advanced statistical software, much of which is open source and credited (e.g., GeoDa in Lersch & Christy, 2020), contemporary analysis has accounted for the spatial structure of data. This can come in the form of spatial lag extensions to traditional OLS models (e.g., Holm & Monaghan, 2021) but increasingly the usage of space–time models based on Bayesian approaches (Ellison et al., 2021; Marco et al., 2017; Quick, 2019).

Langton *et al. Crime Science* (2024) 13:1

Page 10 of 16
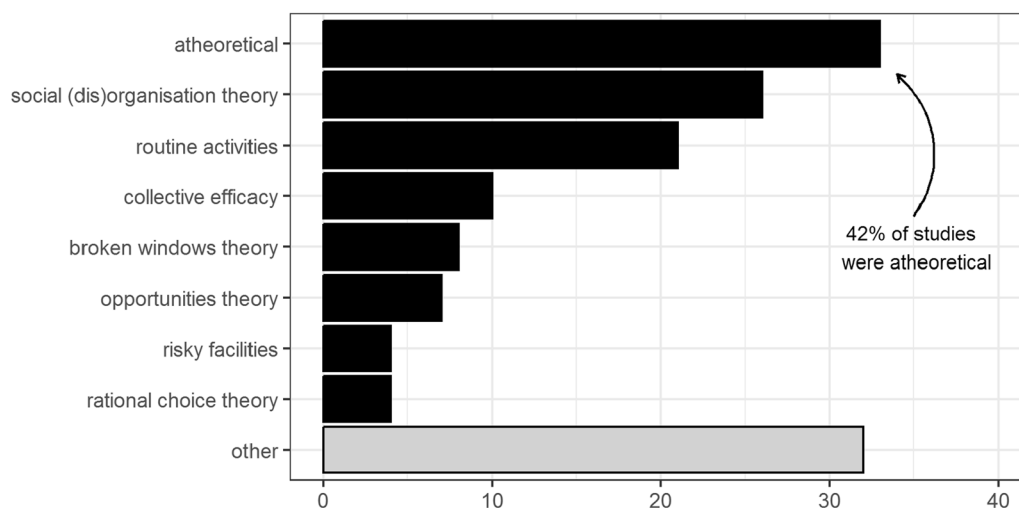
Fig. 6 Frequency of theoretical frameworks used to inform analyses. Note that many studies drew upon more than one theory

We observed that, in our literature sample, the majority of studies employ a deductive approach to running statistical models: there is often an a priori reason to test the association between X and the calls for service type Y. The predictor variables are therefore stated as such, included in a model and results are discussed with reference to their statistical significance at a given threshold. A number of studies used an inductive approach: theory-testing is the goal but the models themselves used an automated procedure of removing variables with large p-values (e.g., Vaughan et al., 2018). Irrespective of the deductive or inductive approach: none of the studies in the literature sample preregistered the research design and hypotheses.

The diversity in the predictor variables used in models is considerable. A total of 473 predictor variable measures thought to explain the spatial patterning of calls for service were used across the whole literature sample. Even after creating an 'other' category for those variables used five or less times, and simplifying individual measures into broad constructs (e.g., deprivation), twenty categories remained. In an effort to summarize the explanatory models used in the literature sample, we traced the predictor variables used for each study, for each call type. By way of demonstration, we plot the explanatory models for mental health-related calls for service in Fig. 7. The equivalent diagram for every call type is available in the Additional file (https://osf.io/jtznc/).

In the cases of mental health and vulnerable people, and indeed across many call types, we see resident-based measures dominating analyses: measures for deprivation, ethnic composition, immigrants, education, family disruption and resident males, among others. These are measures typically derived from social disorganization theory, a theory originally used to explain neighborhood-level concentrations of criminality (as measured using offender residences). Some studies propose that its framework remains relevant for non-crime forms of demand such as mental health (e.g., Vaughan et al., 2018) while others draw upon related explanations via collective efficacy (e.g., White et al., 2019) or community disorder leading to higher exposure to stress (e.g., Lersch et al., 2015).

In addition to resident-based measures, often in the same models, studies have tested the association between risky facilities and (mental health) calls for service. Here, the predictor variables do not capture resident characteristics, but rather, they measure the presence (or volume) of facilities thought to determine the interplay between agents in the ambient population, typically justified through the routine activities framework. Interestingly, the interest in mental health brings a new dimension to this framework. For crime-related calls for service, such as violence, facilities like alcohol outlets increase the risk of victimization through, amongst other things, their role as crime generators and attractors (Kim & Kim, 2022). But for mental health calls, for instance those involving suicide (Lersch & Christy, 2020), there is no offender-target-victim triangle. Instead, risky facilities such as gun shops (coded as 'points of interest') are thought to increase the availability of firearms (which are then used self-inflicted for suicide) while alcohol outlets are considered risk factors due to the problems ensuing from drinking alcohol, such as a reduction in self-control.

Another feature highlighted in Fig. 7, and other call types, is that in our literature sample, studies use multiple
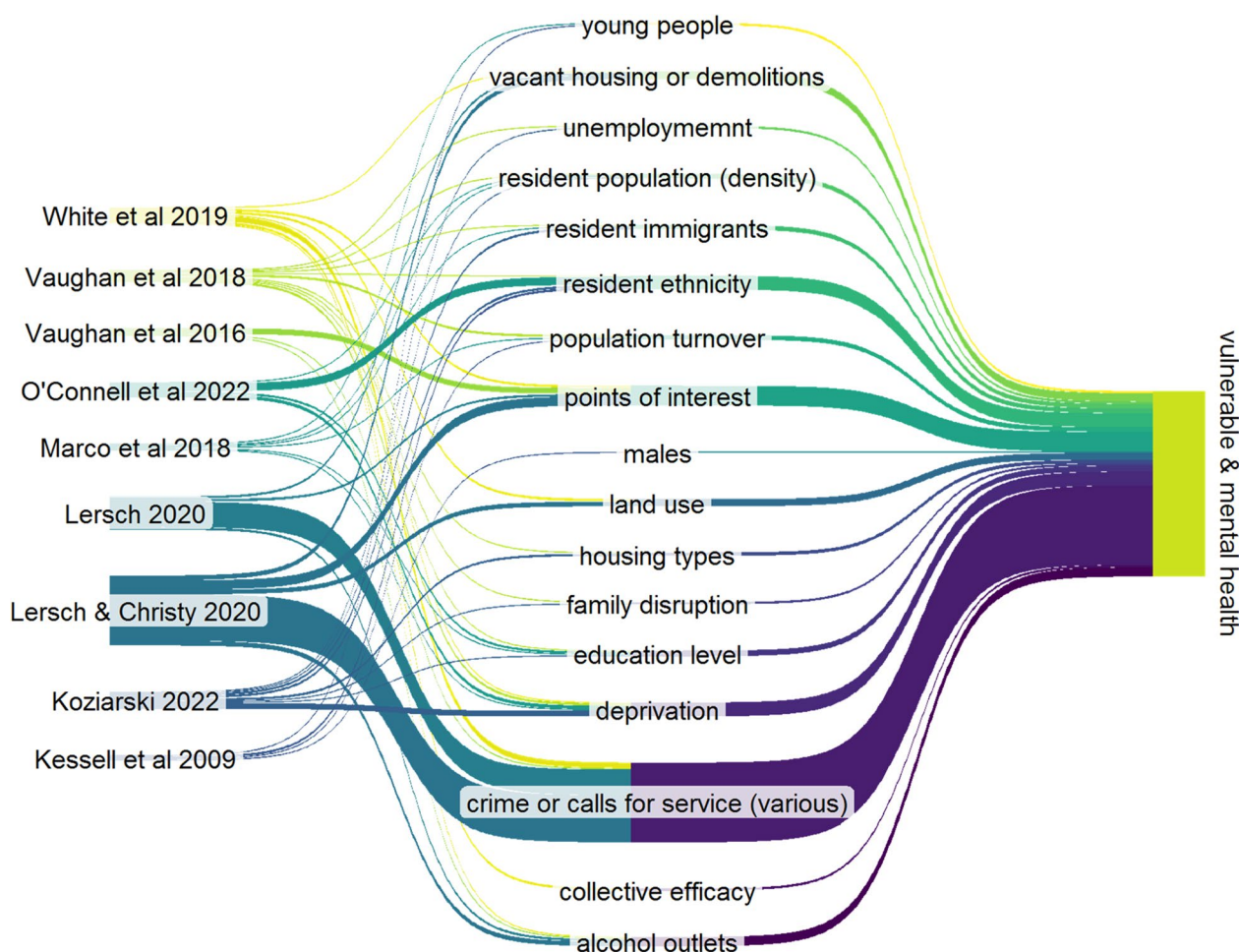
Fig. 7 Study-level summary of explanatory models for vulnerable or mental health calls. All demand types diagrams are available in Additional file (https://osf.io/jtznc/)

measures (and therefore, conduct multiple tests) for the same broad classification of predictor variable. For instance, in identifying the main risk factors for mental health, Lersch and Christy (2020) include numerous different call for service and crime categories that we have then reclassified under 'crime and call for service (various)'. This skew is evident in Fig. 7. It can occur in other ways, for example, not only can multiple measures for deprivation (e.g., average family income, average dwelling value, dwelling in need of major repair, and percentage spending on shelter) be used in the same models, but tests can be repeated for models with a different denominator in the outcome variable [e.g., Andresen and Brantingham (2007)]. We highlight this partly because it has a follow-on effect on our summary of explanatory findings (see next section), but also as a discussion point for future research in terms of the consistency with which broad theoretical concepts, such as deprivation, are tested.

Findings

To summarize the correlational findings, Table 2 presents an overview of the main effect directions reported for multivariable analysis in our literature sample. Here, we only include studies which use a outcome variable that we broadly classify as 'call volume' (e.g., counts, rates) and therefore exclude those with alternative measures which might obscure the summary, such as cluster solutions or the use of Location Quotient.[4] For simplicity, Table 2 aggregates findings together irrespective of denominators in the outcome variable (e.g., ambient populations), but we have flagged such studies for readers interested in the breakdown. Effect directions were simply coded as 'positive', 'negative' or 'none' (i.e., not statistically significant,

---

[4] Nevertheless, studies that used a Location Quotient measure as the outcome in multivariable analysis (Koziarski, 2022; Vaughan et al., 2018) reported very few statistically significant predictors.

Langton *et al. Crime Science*   (2024) 13:1

Page 12 of 16

not a risk as determined by the authors, outside the credibility interval). When appropriate, effect directions were reversed to ensure that the directions reported are consistent with the measure and the respective 'call volume' outcome variable. We choose the single most complete model in each paper to code with the exception of papers which run separate models for different call types: these effect directions are coded separately. A comparable breakdown but for each call type can be reproduced using the data made available in Additional file (https://osf.io/5zshd/). In total, we recorded 831 effect directions. Table 2 summarizes the proportional breakdown of findings, grouped according to the theoretical frameworks that (we deem) are typically used to justify the inclusion of the predictor variable. Within each group, measures are ordered by the frequency of their usage. We have excluded three categories of predictor variable from Table 2, namely, land use, housing types and 'other', because we cannot interpret direction effects. The test-level codings are still available in the corresponding OSF repository.

The only predictor variable for which every model reported a statistically significant finding in the hypothesised direction was resident single persons—a positive association. While the measure was only tested in two studies, the effect was found across multiple call types and in both resident and ambient denominators in the outcome variable (Andresen & Brantingham, 2007; Kim & Kim, 2022). Other measures for a high risk offending resident population, namely, measures of young people and males, performed mixed to poor (28% and 14% positive). Measures of calls for service or crime were tested 89 times and 84% (N = 75) of those found a positive association, confirming expectations that crime and/or high police demand areas tend to co-exist spatially. The three key measures for social disorganization, the most commonly tested theory in our literature sample, have mixed findings on aggregate. Only 31% of studies reported a positive association for deprivation, 9% for ethnic diversity and 33% for population turnover. Of those predictor measures we classify under social disorganization, only physical disorder (67% positive) and unemployment (77% positive) approach a consistent finding.

A contemporary extension of social disorganization, collective efficacy, performed much more in alignment with expectations, but only when a bespoke (survey-based) measure was used. Using such a measure, two thirds of tests found the expected negative association and no study reported the opposite effect. This contrasts considerably with community institutions, often used as a proxy, for which nearly 90% of tests returned no association, the worst performing measure in that regard.

This indicates that its usage as a proxy for collective efficacy, or indeed any measure relevant to emergency police demand, is poor. Alcohol outlets rarely had a negative association with demand (8%) while 46% of tests reported the expected positive association. Other points of interest, which we have aggregated together but tend to be theorized as crime generators or attractors, held a positive association in 50% of cases.

Discussion

This review has provided a descriptive account of studies which have examined the spatial patterning of emergency reactive police demand (ERDP) as typically measured by calls for service data. Our literature sample was obtained via advanced search terms in literature databases (e.g., Web of Science) and forward and backward searches, which resulted in a final sample of 79 studies published between 1989 and 2022. Studies had a variety of motivations for studying the spatial patterning of ERPD, from testing the explanatory power of (combinations of) theoretical frameworks, to examining the impact of different spatial aggregations, exploring emerging demand problems such as mental health crises, or extending existing studies with new methods or data.

The majority of study regions in our sample were in North America, and of those, around two thirds were in the United States. Only one study region lay outside of Europe, North America or Australia: one study conducted in South Korea. This points to a clear skew in the evidence-base. If we as a field are conducting investigations into the spatial patterning of calls for service with the aim of testing 'global' theoretical constructs and proposing laws based on repeated observations of empirical regularities, then we should be precise about the scope of our research: the limits of the samples used and the target population for generalization (Verlaan & Langton, 2023). Here, we can better exploit data availability. Most research has been conducted in the US, for which the data on calls for service, geometry of spatial units of analysis and predictor variables (e.g., census data, Open Street Map) are often openly available. And yet, the body of literature consists of case studies, each of which have rarely been operationalized in a comparable manner, or at least, not comparable *enough* that discrepancies in findings could not potentially be explained by the operationalization alone (e.g., different measures for the same underlying concept, such as deprivation). If the data exist across open data portals for a concerted, uniform examination of the spatial patterning of ERPD (e.g., its concentration, correlates), we would encourage such an endeavour. In the age of widespread data availability and (open) software for computational analysis, there is no reason why analyses cannot be scaled-up beyond single

case studies.[5] Without this, and instead with a continued reliance on one-off (often uniquely operationalized) case studies, there is a risk that we fail to identify empirical regularities, or take too long to dismiss theoretical expectations which are misguided.

There is a general preference for studies in our sample to examine crime-related calls for service. The most common call types used, by some margin, were those involving violence or assault. Studies tended not to provide specific justification for their call type choice, but we would speculate that this preference reflects that violence is a high-harm (societally relevant) and relatively common (societally relevant, data sufficiency at small aggregations) crime type. The most common non-crime call type were those involving mental health or vulnerable people. The vast majority of studies examining this call type were in recent years (post-2016), reflecting growing concerns over the (increasing) public reliance on the police for mental health support (CoP, 2015) and the recognition that a considerable proportion of dispatched police deployment time involves health crises (Ratcliffe, 2021).

The selection (and justification) for a call type will ultimately be determined by the goals of the study, but we would encourage future research to proactively engage with why call type(s) are selected and relatedly how they are defined. Defining incidents as per the dispatch center categorization is a non-trivial decision: can domestic and non-domestic violent incidents be disentangled? This would impact on, for example, the relevance of resident-based predictor variables in explaining spatial patterns. Does theft include a flag for indoor or outdoor incidents? Routine activity-inspired predictors, such as the presence of risky facilities (e.g., transport stations), might only have the expected association with such a distinction. As noted above with regards the scalability, the data the software are accessible to analyse multiple (even, all) call types simultaneously. This offers opportunity for the field to move away from one-off, heterogeneously operationalized case studies, and towards a comprehensive understanding of the concentration and correlates of different call types.

Despite calls for service data often being recorded and made available at point or street-level, and the demonstrable benefits of such fine-grained scales (Steenbeek & Weisburd, 2016), there is still a preference for meso-level analysis. We would attribute this largely to the popularity of studies drawing upon social disorganization theory: a framework originally thought to manifest at the neighborhood-level, and one requiring resident-based characteristics to measure, which tend to only be available using census (neighborhood) units. In our sample, proposals for social disorganization to be tested using micro-level units of analysis remain unrealised (Weisburd et al., 2012). By contrast, univariate analysis (e.g., descriptions of concentration) can more often be conducted at the micro-level, as can correlational analyses of opportunity theories (as is typically theorized), due to data availability. Studies that test social disorganization and opportunity theoretical frameworks simultaneously might then aggregate data to the larger of the two spatial scales, masking detail and diluting the theoretical validity of predictor measures thought to operate at the micro-level. Here, we might make better usage of multilevel frameworks to test theories that operate at different spatial scales, and the cross-level interaction between them (e.g., micro-level correlations being stronger within certain neighborhoods).

Our summary of correlational findings in the literature sample were, on the whole, mixed. We would propose that there is only one clear empirical regularity in our literature sample: the co-occurrence of different call and crime types in space. Irrespective of the spatial scale used, different types of emergency calls for police service and/or crime tend to concentrate in the same geographic spaces. As advised above, by exploiting open data and open software capable of computational analysis, we could further unpick these commonalities to identify which call types tend to co-exist, and to what extent these similarities exist across different contexts. This would point towards common underlying demand-generating mechanisms and help guide the selection (and combinations) of call types investigated in subsequent research. A number of other predictor variables showed some consistency in correlational findings, including collective efficacy (negative), resident single persons (positive), unemployment (positive), and physical disorder (positive). Perhaps the most troubling finding, given the popularity and longevity of the theory, was that the three variables most commonly used to measure social disorganization (deprivation, population turnover, ethnic diversity) often performed poorly. Given these mixed findings, future research might find the trade-off required to test such associations, namely, aggregating data to meso-level scales, difficult to justify, unless there are specific reasons for using resident-based measures (e.g., investigating a resident-specific call type).

This study is not without caveats. First, our search terms focused on variants of 'calls for service' and 'demand' in titles and abstracts, with the meso and micro-level criteria being applied in subsequent screening. This

---

[5] One exemplar study is Bennett (2018)'s paper on the predictors of police response times across 40 police agencies in the US. The author obtained their data via Freedom of Information requests. But, in our case, without the need for information on response times, geocoded CAD data from open data portals are sufficient and widely available across many police jurisdictions (see for example the summary by the Police Data Initiative: https://www.policedatainitiative.org/datasets/calls-for-service/. Last accessed 28 November 2023).

approach is unlikely to capture studies that, in examining the 'spatial patterning of [crime type]', clarify their use of calls for service data (as a measure for a crime) in the paper itself. Another approach would have been to search for spatial terms in combination with crime or policing, then subsequently narrow by calls for service. Second, given the wide variety of associations tested in the literature, we had to simplify in order to make a summary of findings. This is in part a consequence of conducting a scoping review rather than a systematic review (with a critical appraisal) that could have permitted more detailed analysis of a specific question, such as the effectiveness of a policing intervention. We focus exclusively on correlational findings (rather than, for instance, experimental studies) which do not have strong claims for causality, but rather, describe the typical characteristics of high-demand areas. We also only categorize associations according to 'positive', 'negative' or 'none' without a further classification of effect size or scrutiny of the study's strengths (or shortcomings). Nevertheless, correlational findings form the evidence-base from which empirical regularities are observed and then subsequently investigated. So, we expect our findings to provide a useful guide for future research in this regard. We would encourage others to make use of our open materials for the purposes of extending the findings presented here.

Conclusion

This review has provided a descriptive account of studies which have sought to describe and/or explain the spatial patterning of emergency reactive police demand ('ERPD'), as measured using calls for service data. The study had three principal aims, namely, to synthesize correlational findings for the purposes of establishing 'empirical regularities', to gauge the extent to which different forms of (crime and non-crime) police demand are being examined, and to ascertain whether the field has embraced the usage of fine-grained, 'micro' spatial scales, such as street segments. We found minimal evidence for empirical regularities, even among common predictors such as deprivation. Resident single persons, unemployment, physical disorder, crime/other calls for service, and collective efficacy tended to have the hypothesized associations. The focus of most analyses in the field is on calls for service related to crime, such as violence, although there has been substantial attention paid to calls involving mental health and vulnerable people in recent years. Despite recent calls in crime and place research to use micro-level spatial scales, there is still a preference to study ERPD using meso-level units such as census blocks. One key area in which we identify shortcomings is the reliance on one-off case studies, and consequently, the inconsistency with which analyses are conducted and theories tested. We propose that the field better exploits the widespread availability of relevant open data (in the US) and open software. Without doing so, our evidence-base suffers: generalizability is limited and empirical regularities remain undiscovered.

Appendix
Primary database searches
See Table 3.

Langton *et al. Crime Science*   (2024) 13:1

Page 15 of 16

Table 3 Search terms within on each database. Queries were designed to be comparable across databases. Total includes duplicates

Database	Search	Search terms	Results
Scopus	#1	TITLE-ABS-KEY (police OR policing OR "law enforcement")	114,266
	#2	TITLE-ABS-KEY ((call* AND (service OR assistance OR emergenc*)) OR "call* W/2 police" OR ((emergenc* AND (reactive OR public OR citizen*) AND demand*)))	167,361
	#3	#1 AND #2	2397
Web of Science	#1	TS = (police OR policing OR "law enforcement")	68,378
	#2	TS = ((call* AND (service OR assistance OR emergenc*)) OR "call* NEAR/2 police" OR ((emergenc* AND (reactive OR public OR citizen*) AND demand*)))	70,047
	#3	#1 AND #2	1226
Criminal Justice Abstracts	#1	TI (police OR policing OR "law enforcement") OR AB (police OR policing OR "law enforcement") OR KW (police OR policing OR "law enforcement")	52,558
	#2	TI ((call* AND (service OR assistance OR emergenc*)) OR "call* N2 police" OR ((emergenc* AND (reactive OR public OR citizen*) AND demand*))) OR AB ((call* AND (service OR assistance OR emergenc*)) OR "call* N2 police" OR ((emergenc* AND (reactive OR public OR citizen*) AND demand*))) OR KW ((call* AND (service OR assistance OR emergenc*)) OR "call* N2 police" OR ((emergenc* AND (reactive OR public OR citizen*) AND demand*)))	3115
	#3	#1 AND #2	780
IBSS	#1	MAINSUBJECT.EXACT("Police" OR "Law enforcement") OR TI(police OR policing OR "law enforcement") OR AB(police OR policing OR "law enforcement")	40,188
	#2	TI((call* AND (service OR assistance OR emergenc*)) OR "call* NEAR/2 police" OR ((emergenc* AND (reactive OR public OR citizen*) AND demand*))) OR AB((call* AND (service OR assistance OR emergenc*)) OR "call* NEAR/2 police" OR ((emergenc* AND (reactive OR public OR citizen*) AND demand*)))	13,687
	#3	#1 AND #2	511
Total			4914

References
Andresen, M. A. (2006a). A spatial analysis of crime in Vancouver, British Columbia: A synthesis of social disorganization and routine activ-ity theory. *The Canadian Geographer/le Géographe Canadien, 50*(4), 487–502.
Andresen, M. A. (2006b). Crime measures and the spatial analysis of criminal activity. *British Journal of Criminology, 46*(2), 258–285.

Andresen, M. A. (2011a). Estimating the probability of local crime clusters: The impact of immediate spatial neighbors. *Journal of Criminal Justice, 39*(5), 394–404.
Andresen, M. A. (2011b). The ambient population and crime analysis. *The Professional Geographer, 63*(2), 193–212.
Andresen, M. A., & Brantingham, P. J. (2007). *Hot spots of crime in Vancouver and their relationship with population characteristics*. Department of Justice Canada.
Andresen, M. A., & Linning, S. J. (2012). The (in) appropriateness of aggregat-ing across crime types. *Applied Geography, 35*(1–2), 275–282.
Andresen, M. A., & Malleson, N. (2011). Testing the stability of crime patterns: Implications for theory and policy. *Journal of Research in Crime and Delinquency, 48*(1), 58–82.
Bennett, D. S. (2018). *Police response times to calls for service*.
Bocker Parks, R. (2015). *Examining spatiotemporal patterns of disorder at bars in Newark, NJ* [PhD thesis]. Rutgers University-Graduate School-Newark.
Boggess, L. N., & Maskaly, J. (2014). The spatial context of the disorder–crime relationship in a study of Reno neighborhoods. *Social Science Research, 43*, 168–183.
Boulton, L., McManus, M., Metcalfe, L., Brian, D., & Dawson, I. (2017). Calls for police service: Understanding the demand profile and the UK police response. *The Police Journal, 90*(1), 70–85.
Braga, A. A., Turchan, B. S., Papachristos, A. V., & Hureau, D. M. (2019). Hot spots policing and crime reduction: An update of an ongoing sys-tematic review and meta-analysis. *Journal of Experimental Criminology, 15*(3), 289–311.
Chohlas-Wood, A., Merali, A., Reed, W., & Damoulas, T. (2015). Mining 911 calls in New York City: Temporal patterns, detection, and forecasting. *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*.
Clare, J., Townsley, M., Birks, D. J., & Garis, L. (2019). *Policing A Journal of Policy and Practice, 13*(3), 286–299.
CoP. (2015). College of policing analysis: Estimating demand on the police service. In *College of Policing*. https://paas-s3-broker-prod-lon-6453d 964-1d1a-432a-9260-5e0ba7d2fc51.s3.eu-west-2.amazonaws.com/ s3fs-public/2021-03/demand-on-policing-report.pdf

Dewinter, M., Vandeviver, C., Dau, P. M., Vander Beken, T., & Witlox, F. (2022). Hot spots and burning times: A spatiotemporal analysis of calls for service to establish police demand. *Applied Geography, 143*, 102712.

Ellison, M., Bannister, J., Lee, W. D., & Haleem, M. S. (2021). Understanding policing demand and deployment through the lens of the city and with the application of big data. *Urban Studies, 58*, 3157–3175.

Guidi, S., Homel, R., & Townsley, M. (1997). *Hot spots and repeat break and enter crimes: An analysis of police calls for service data*. Research; Coordination Division, Criminal Justice Commission.

Haddaway, N. R., Collins, A. M., Coughlin, D., & Kirk, S. (2015). The role of google scholar in evidence reviews and its applicability to grey literature searching. *PLoS ONE, 10*(9), e0138237.

Hagan, J., McCarthy, B., Herda, D., & Cann Chandrasekher, A. (2018). Dual-process theory of racial isolation, legal cynicism, and reported crime. *Proceedings of the National Academy of Sciences, 115*(28), 7190–7199.

Harzing, A.-W. (2010). *The publish or perish book*. Tarma Software Research Pty Limited Melbourne.

Hibdon, J., Telep, C. W., & Groff, E. R. (2017). The concentration and stability of drug activity in Seattle, Washington using police and emergency medical services data. *Journal of Quantitative Criminology, 33*, 497–517.

Hodgkinson, T., & Andresen, M. A. (2019). Understanding the spatial patterns of police activity and mental health in a Canadian city. *Journal of Contemporary Criminal Justice, 35*(2), 221–240.

Hodgkinson, T., Saville, G., & Andresen, M. A. (2020). The diffusion of detriment: Tracking displacement using a city-wide mixed methods approach. *The British Journal of Criminology, 60*(1), 198–218.

Jones, B., Tiwari, C., & Ramisetty-Mikler, S. (2019). A data-driven approach using 911-call data to reconfigure police beat zones. *Papers in Applied Geography, 5*(1–2), 166–175.

Kim, Y.-A., & Kim, J. (2022). Examining the effects of physical environment and structural characteristics on the spatial patterns of crime in Daegu, South Korea. *Crime & Delinquency*, 00111287221130963.

Koziarski, J. (2021). Examining the spatial concentration of mental health calls for police service in a small city. *Policing A Journal of Policy and Practice, 15*(2), 1011–1028.

Koziarski, J. (2022). *The spatial concentration, stability, and specialization of mental health calls for service: Evidence in support of proactive, place-based interventions* [PhD thesis]. The University of Western Ontario (Canada).

Langton, S., Ruiter, S., & Verlaan, T. (2022). Describing the scale and composition of calls for police service: A replication and extension using open data. *Police Practice and Research*, 1–16.

Laufs, J., Bowers, K., Birks, D., & Johnson, S. D. (2021). Understanding the concept of 'demand' in policing: A scoping review and resulting implications for demand management. *Policing and Society, 31*(8), 895–918.

Lersch, K., & Christy, A. (2020). The geography of mental health: An examination of police calls for service. *Geographies of Behavioural Health, Crime, and Disorder: The Intersection of Social Problems and Place*, 13–36.

Lersch, K., Sellers, C., & Cromwell, P. (2015). Residential foreclosure rates and calls for service for domestic disputes: An exploratory analysis. *American Journal of Criminal Justice, 40*, 579–592.

Louis, S., & Greene, J. R. (2020). Social context in police legitimacy: Giving meaning to police/community contacts. *Policing and Society, 30*(6), 656–673.

Luan, H., Quick, M., & Law, J. (2016). Analyzing local spatio-temporal patterns of police calls-for-service using Bayesian integrated nested laplace approximation. *ISPRS International Journal of Geo-Information, 5*(9), 162.

Lymperopoulou, K., Bannister, J., & Krzemieniewska-Nandwani, K. (2022). Inequality in exposure to crime, social disorganization and collective efficacy: Evidence from greater manchester, united kingdom. *The British Journal of Criminology, 62*(4), 1019–1035.

Marco, M., Gracia, E., Lopez-Quilez, A., & Lila, M. (2018). What calls for service tell us about suicide: A 7-year spatio-temporal analysis of neighborhood correlates of suicide-related calls. *Scientific Reports, 8*(1), 6746.

Marco, M., Lpez-Qulez, A., Conesa, D., Gracia, E., & Lila, M. (2017). Spatio-temporal analysis of suicide-related emergency calls. *International Journal of Environmental Research and Public Health, 14*(7), 735.

Pfeiffer, D., & Lucio, J. (2015). An unexpected geography of opportunity in the wake of the foreclosure crisis: Low-income renters in investor-purchased foreclosures in phoenix, Arizona. *Urban Geography, 36*(8), 1197–1220.

Pierce, G. L., Spaar, S., & Briggs, L. R. (1988). *The character of police work: Strategic and tactical implications*. Center for Applied Social Research, Northeastern University.

Quick, M. (2019). *Spatiotemporal crime patterns and the urban environment: Evidence for planning and place-based policy*.

R Core Team. (2022). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ratcliffe, J. H. (2021). Policing and public health calls for service in philadelphia. *Crime Science, 10*(1), 1–6.

Reinhard, D. (2023). The influence of vacation home rentals on neighborhood crime and disorder. *American Journal of Criminal Justice, 48*(1), 233–249.

Roh, S., & Choo, T. M. (2008). Looking inside zone v: Testing social disorganization theory in suburban areas. *W. Criminology Rev., 9*, 1.

Roman, C. G., & Reid, S. E. (2012). Assessing the relationship between alcohol outlets and domestic violence: Routine activities and the neighborhood environment. *Violence and Victims, 27*(5), 811–828.

Sherman, L. W., Gartin, P. R., & Buerger, M. E. (1989). Hot spots of predatory crime: Routine activities and the criminology of place. *Criminology, 27*(1), 27–56.

Spelman, W. (1995). Criminal careers of public places. *Crime and Place, 4*, 115–144.

Steenbeek, W., & Weisburd, D. (2016). Where the action is in crime? An examination of variability of crime across different spatial units in the Hague, 2001–2009. *Journal of Quantitative Criminology, 32*(3), 449–469.

Taniguchi, T. A., & Salvatore, C. (2018). Policing a negotiated world: A partial test of Klinger's ecological theory of policing. *Journal of Quantitative Criminology, 34*, 345–366.

van de Schoot, R., de Bruin, J., Schram, R., Zahedi, P., de Boer, J., Weijdema, F., Kramer, B., Huijts, M., Hoogerwerf, M., Ferdinands, G., et al. (2021). An open source machine learning framework for efficient and transparent systematic reviews. *Nature Machine Intelligence, 3*(2), 125–133.

van Holm, E. J., & Monaghan, J. (2021). The relationship of Airbnb to neighborhood calls for service in three cities. *Cities, 116*, 103241.

Vaughan, A. D., Hewitt, A. N., Andresen, M. A., & Brantingham, P. L. (2016). Exploring the role of the environmental context in the spatial distribution of calls-for-service associated with emotionally disturbed persons. *Policing A Journal of Policy and Practice, 10*(2), 121–133.

Vaughan, A. D., Ly, M., Andresen, M. A., Wuschke, K., Hodgkinson, T., & Campbell, A. (2018). Concentrations and specialization of mental health–related calls for police service. *Victims & Offenders, 13*(8), 1153–1170.

Verlaan, T., & Langton, S. (2023). On the use of inferential statistics on administrative police data. In L. Huey & D. Buil-Gil (Eds.), *The crime data handbook*. Bristol University Press.

Warner, B. D., & Pierce, G. L. (1993). Reexamining social disorganization theory using calls to the police as a measure of crime. *Criminology, 31*(4), 493–517.

Weisburd, D. (2015). The law of crime concentration and the criminology of place. *Criminology, 53*(2), 133–157.

Weisburd, D., Bernasco, W., & Bruinsma, G. (2008). *Putting crime in its place*. Springer.

Weisburd, D., Groff, E. R., & Yang, S.-M. (2012). *The criminology of place: Street segments and our understanding of the crime problem*. Oxford University Press.

White, C., Goldberg, V., Hibdon, J., & Weisburd, D. (2019). Understanding the role of service providers, land use, and resident characteristics on the occurrence of mental health crisis calls to the police. *Journal of Community Psychology, 47*(8), 1961–1982.

Publisher's Note