

PEMETAAN SPASIO-TEMPORAL DAN PREDIKSI PENYAKIT PRIORITAS BERBASIS DATA RUTIN PUSKESMAS DARATAN DAN KEPULAUAN DI KABUPATEN PANGKEP

Idarwati Mustafa^{*1}, Andi Mikhrunnisai², A. Nur Azizah³, Andi Sada Potto⁴

^{1,2,3}Program Studi Administrasi Kesehatan, Universitas Puangrimaggalatung,
Sengkang, Indonesia

⁴Dinas Kesehatan Kabupaten Pangkajene dan Kepulauan, Pangkajene, Indonesia

* Corresponding Author: idarwati.mustafa@gmail.com

ARTICLE INFO

Article history:

Received 19-10-2026

Revised 20-10-2026

Accepted 21-10-2026

Available online 24-10-2026

Kata Kunci:

Data rutin puskesmas; Spasio-temporal; Wilayah kepulauan; Machine learning; Perencanaan kesehatan.

Keywords:

Routine health data; Spatio-temporal; Island region; Machine learning; Health planning.

ABSTRAK

Data rutin puskesmas di Kabupaten Pangkajene dan Kepulauan masih lebih banyak dipakai untuk pelaporan administratif daripada perencanaan program berbasis wilayah, padahal wilayah daratan dan kepulauan memiliki masalah kesehatan yang berbeda. Penelitian ini bertujuan untuk memetakan pola spasial dan temporal lima kelompok penyakit prioritas serta mengevaluasi kemampuan model *machine learning* dalam memprediksi jumlah kasus baru bulanan di tingkat desa. Penelitian kuantitatif retrospektif ini menggunakan kerangka *Cross-Industry Standard Process for Data Mining* dari dua puskesmas daratan dan dua puskesmas kepulauan yang mencakup 21 desa selama Januari 2024 hingga Juni 2026. Analisis dilakukan menggunakan angka kasus baru per 1.000 penduduk, *rate ratio*, regresi Poisson dengan

komponen musiman, *standardized incidence ratio* yang dihaluskan dengan *empirical Bayes*, *Moran's I*, serta perbandingan antara *random forest*, *gradient boosting*, dan model pembandingan sederhana. Selama periode penelitian tercatat 176.561 kunjungan, dengan 149.042 di antaranya merupakan kasus baru. Angka hipertensi di desa kepulauan 3,76 kali lebih tinggi daripada desa daratan dan membentuk pengelompokan spasial, sedangkan infeksi saluran pernapasan atas akut dan diare lebih banyak ditemukan di daratan. Kasus hipertensi meningkat sekitar 28,6% setiap tahun. Model *random forest* menghasilkan *mean absolute error* sebesar 4,88 kasus, sedikit lebih baik daripada *naive model* dan setara dengan *moving average* tiga bulan. Perencanaan program perlu berbasis desa, dan model prediksi sederhana sudah cukup untuk mendukung *dashboard* perencanaan puskesmas.

ABSTRACT

Routine primary health care data in Pangkajene and Islands Regency are still used mainly for administrative reporting, although mainland and island areas face different health problems. This study aimed to map the spatial and temporal patterns of five priority disease groups and to evaluate how well machine learning models predict monthly new cases at village level. This retrospective quantitative study followed the *Cross-Industry Standard Process for Data Mining* from two mainland and two island primary health centres serving 21 villages from January 2024 to June 2026. The analysis used new-case rates per 1,000 population, rate ratios, Poisson regression with a seasonal component, standardised incidence ratios smoothed with empirical Bayes, *Moran's I*, and a comparison of random forest, gradient boosting and simple benchmark models. Over the study period, 176,561 visits were recorded, of which 149,042 were new cases. The hypertension rate in island

villages was 3.76 times higher than in mainland villages and showed spatial clustering, while acute upper respiratory infection and diarrhoea were more common on the mainland. Hypertension cases rose by about 28.6% per year. Random forest produced a mean absolute error of 4.88 cases, slightly better than the naive model and similar to a three-month moving average. Program planning should be village-based, and simple forecasting models are sufficient to support a primary care planning dashboard.

This is an open access article under the [CC BY-NC](#) license.

Copyright © 2021 by Author. Published by Universitas Bina Bangsa Getsempena



PENDAHULUAN

Setiap hari puskesmas mencatat data kunjungan pasien, mulai dari registrasi, diagnosis yang dikode dengan ICD-10, sampai status kasus baru atau lama. Apabila dikelola dengan baik, data tersebut dapat menggambarkan beban penyakit menurut wilayah dan waktu. Kenyataannya, beberapa evaluasi terhadap sistem informasi manajemen puskesmas (SIMPUS) di Indonesia menunjukkan bahwa data tersebut masih lebih banyak digunakan untuk pencatatan dan pelaporan. Di Kota Yogyakarta, misalnya, pengumpulan data sebagian besar masih dilakukan secara manual, laporan belum tepat waktu dan belum lengkap, serta informasi yang dihasilkan jarang digunakan dalam pengambilan keputusan (Yani & Lazuardi, 2018). Evaluasi menggunakan model HOT-fit di Kabupaten Purworejo juga menemukan bahwa sistem yang mudah digunakan belum tentu menghasilkan data yang lengkap, karena layanan di puskesmas belum tercatat, sementara prosedur operasional, pelatihan, dan monitoring rutin belum tersedia (Thenu et al., 2016). Selain itu, pemanfaatan SIMPUS oleh petugas dipengaruhi oleh faktor informasi, ekonomi, kontrol, dan layanan (Rahmi et al., 2024).

Masalah kualitas data tidak hanya berkaitan dengan kelengkapan laporan, tetapi juga dengan ketepatan kode diagnosis. Penelitian di Puskesmas Taram menemukan bahwa 72,9% kode ICD-10 diabetes melitus pada rekam medis rawat jalan tidak tepat, terutama karena penulisan diagnosis yang tidak jelas dan tidak lengkap (Ornela et al., 2025). Kesalahan pengodean semacam ini dapat menyebabkan bias dalam perencanaan kesehatan, karena penyakit yang tampak dominan dalam laporan belum tentu mencerminkan kondisi epidemiologis yang sebenarnya, melainkan dapat dipengaruhi oleh praktik pengodean yang diterapkan di puskesmas tertentu.

Kondisi geografis Kabupaten Pangkajene dan Kepulauan (Pangkep) menambah kompleksitas persoalan tersebut. Sebagian wilayahnya berupa gugusan pulau kecil yang hanya dapat dicapai dengan perahu, sehingga akses layanan dan alur pelaporan data berbeda dengan daratan. Di daerah terpencil, hambatan geografis, hambatan fisik, dan ketersediaan tenaga kesehatan terbukti berhubungan dengan akses pelayanan (Nurfadillah

et al., 2025). Jarak tempuh, cuaca yang sulit diprediksi, dan biaya transportasi juga membatasi kunjungan masyarakat ke puskesmas (Siregar et al., 2025). Sejak tahun 2017, Puskesmas Sabutung mengembangkan layanan kesehatan dasar bergerak bernama Perahu Sehat Pulau Bahagia untuk menjangkau masyarakat di pulau-pulau, dan model ini kemudian diadaptasi oleh beberapa puskesmas lain (Budiman & Aruan, 2021). Di sisi lain, Rencana Strategis Dinas Kesehatan Pangkep mencatat bahwa data antarprogram di puskesmas belum sinkron sehingga menyulitkan proses perencanaan (Dinas Kesehatan Kabupaten Pangkajene dan Kepulauan, 2025).

Pendekatan spasial sudah cukup banyak digunakan dalam penelitian kesehatan di Indonesia. Analisis autokorelasi spasial menunjukkan bahwa kasus demam berdarah *dengue* cenderung mengelompok di lokasi tertentu dalam wilayah kerja puskesmas (Syamsir et al., 2020) dan berkaitan dengan kepadatan penduduk, angka bebas jentik, serta curah hujan di tingkat desa (Hidayati, 2023). Pemanfaatan sistem informasi geografis juga membantu rumah sakit mengetahui asal pasien dan pola kunjungan sebagai dasar perencanaan layanan (Rindiyan et al., 2025). Di dinas kesehatan, *dashboard* telah dikembangkan untuk memantau standar pelayanan minimal dan program kesehatan keluarga (Hizriansyah, 2023; Sari et al., 2022; Putri & Ramani, 2024), termasuk melalui pelatihan bagi pengelola data (Andayani et al., 2025). Untuk keperluan prediksi, *systematic review* yang dilakukan Ajulo et al. (2025) mencatat meningkatnya penggunaan model *machine learning* spasial dan spasio-temporal, tetapi cara pelaporan metodenya sangat beragam dan model jarang dibandingkan dengan pembanding sederhana. Jaya et al. (2025) bahkan menemukan bahwa XGBoost dan *neural network* memberikan hasil yang sangat baik pada data latih, namun kinerjanya menurun tajam ketika diuji pada data baru dalam prakiraan *dengue* di Bandung.

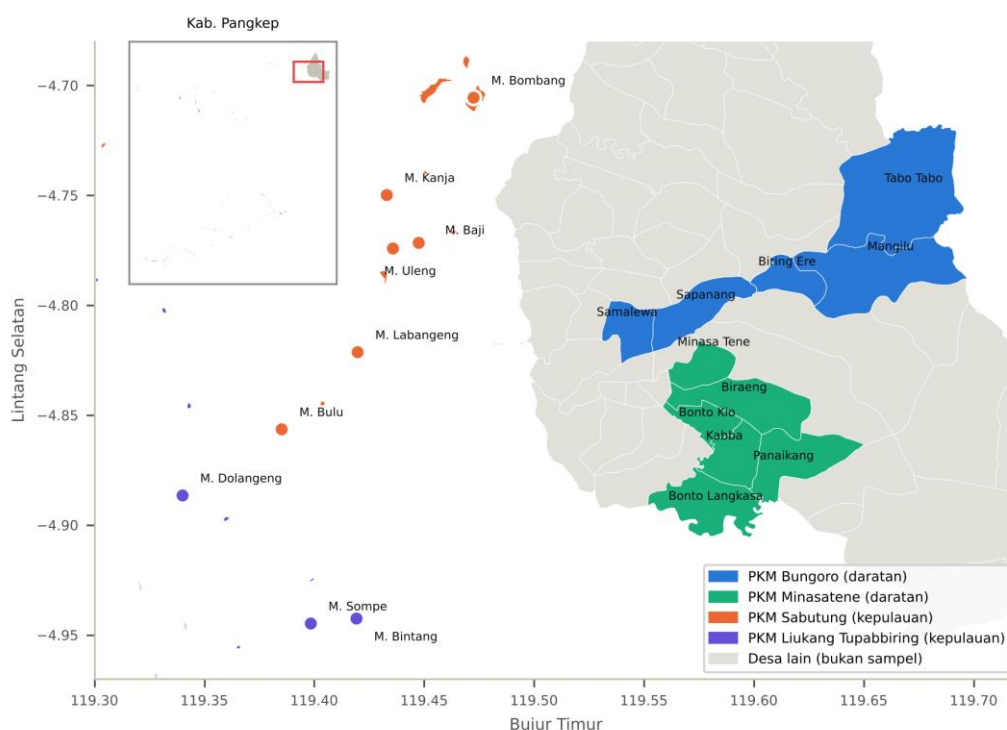
Berdasarkan uraian tersebut, masih terdapat beberapa kekosongan penelitian. Sebagian besar studi spasial di Indonesia hanya membahas satu penyakit menular, terutama *dengue*, sehingga analisis multipenyakit dari data rutin puskesmas, termasuk penyakit tidak menular, masih jarang dilakukan. Perbandingan antara wilayah daratan dan kepulauan dalam satu kabupaten dengan unit analisis desa juga belum banyak dilaporkan. Selain itu, model prediksi berbasis *machine learning* jarang dibandingkan dengan *naive model* yang sebenarnya dapat dihitung sendiri oleh petugas puskesmas. Oleh karena itu, penelitian ini bertujuan untuk (1) menggambarkan karakteristik dan kualitas data rutin empat puskesmas di Kabupaten Pangkep, (2) memetakan pola spasial dan temporal lima kelompok penyakit prioritas di tingkat desa, dan (3) mengevaluasi kinerja model *machine*

learning dalam memprediksi kasus baru bulanan sebagai dasar pengembangan *dashboard* perencanaan program.

METODE PENELITIAN

Desain, Lokasi, dan Sumber Data

Penelitian ini menggunakan desain kuantitatif analitik dengan pendekatan retrospektif. Alur penelitian mengikuti kerangka *Cross-Industry Standard Process for Data Mining* (CRISP-DM) yang terdiri atas enam tahap, yaitu pemahaman masalah, pemahaman data, persiapan data, pemodelan, evaluasi, dan penerapan (Schröer et al., 2021). Penelitian dilaksanakan di empat puskesmas yang dipilih secara *purposive* untuk mewakili perbedaan karakteristik geografis, yaitu Puskesmas Bungoro dan Puskesmas Minasatene di wilayah daratan serta Puskesmas Sabutung dan Puskesmas Liukang Tupabbiring di wilayah kepulauan (Gambar 1). Unit analisis adalah desa atau kelurahan tempat tinggal pasien, yang berjumlah 21 desa, terdiri atas 11 desa daratan dan 10 desa kepulauan.



Gambar 1. Lokasi empat puskesmas dan 21 desa wilayah studi di Kabupaten Pangkajene dan Kepulauan

Sumber data utama adalah rekapitulasi bulanan diagnosis pasien rawat jalan di keempat puskesmas selama Januari 2024 sampai Juni 2026. Setiap baris berisi tahun, bulan, puskesmas, desa dan kode wilayah 10 digit, kategori wilayah, kode ICD-10 tiga karakter, diagnosis, status kasus, dan jumlah kunjungan. Data jumlah penduduk desa tahun 2024 diperoleh dari publikasi Kecamatan Dalam Angka 2025 (Badan Pusat Statistik Kabupaten

Pangkajene dan Kepulauan, 2025), sedangkan persentase penduduk dengan pendidikan tertinggi tamat sekolah dasar diperoleh dari data konsolidasi bersih kependudukan semester 2 tahun 2025 (Dinas Kependudukan dan Pencatatan Sipil Kabupaten Pangkajene dan Kepulauan, 2026). Batas administrasi desa, luas wilayah, persentase penduduk yang bekerja sebagai nelayan, dan persentase penduduk berusia 60 tahun ke atas diambil dari basis data Kementerian Dalam Negeri tahun 2020 (Kementerian Dalam Negeri, 2020). Seluruh data yang digunakan merupakan data agregat dan tidak memuat identitas pasien.

Persiapan Data

Pada tahap persiapan data, peneliti memeriksa kelengkapan data untuk setiap bulan dan desa, mencocokkan kode ICD-10 dengan label diagnosisnya, lalu menggabungkan data kunjungan dengan data penduduk dan peta desa menggunakan kode desa 10 digit. Analisis penyakit hanya menggunakan kunjungan dengan status kasus baru. Kode ICD-10 kemudian dikelompokkan ke dalam lima kelompok penyakit prioritas yang paling sering ditemukan dan berkaitan langsung dengan program puskesmas, yaitu hipertensi (I10-I15), infeksi saluran pernapasan atas akut atau ISPA (J00-J06), gastritis dan dispepsia (K29-K30), diabetes melitus (E10-E14), serta diare dan gastroenteritis (A00-A09). Hasil pengelompokan disusun menjadi data panel desa $\times$ kelompok penyakit $\times$ bulan sebanyak 3.150 observasi, dan kombinasi yang tidak memiliki kunjungan diberi nilai nol.

Analisis Spasio-Temporal

Angka kasus baru dinyatakan per 1.000 penduduk per tahun, yaitu jumlah kasus selama 30 bulan dibagi jumlah penduduk desa tahun 2024 yang dikalikan 2,5 tahun. Tahun 2024 dipilih karena berada di pertengahan periode penelitian, dan analisis sensitivitas memakai data Dukcapil 2025 serta Kementerian Dalam Negeri 2020. Perbedaan angka antara kepulauan dan daratan dinyatakan dalam *rate ratio* (RR) beserta interval kepercayaan 95% berdasarkan distribusi Poisson. Tren dianalisis dengan regresi Poisson log-linear yang memuat variabel waktu dan pasangan fungsi sinus-kosinus berperiode 12 bulan sebagai komponen musiman. *Standard error* dikoreksi terhadap *overdispersion*, kemudian koefisien waktu diubah menjadi persentase perubahan per tahun. Arah tren juga diuji dengan uji Mann-Kendall.

Risiko relatif masing-masing desa dihitung sebagai *standardized incidence ratio* (SIR), yaitu rasio antara jumlah kasus yang teramati dan jumlah kasus yang diharapkan:

$$SIR_i = \frac{O_i}{E_i} = \frac{O_i}{n_i \cdot r} \quad (1)$$

O_i menyatakan jumlah kasus baru yang teramati di desa i , n_i menyatakan jumlah orang-tahun, dan r adalah angka kasus seluruh 21 desa. Jumlah penduduk desa kepulauan relatif kecil sehingga angka kasusnya mudah berfluktuasi. Untuk itu, SIR dihaluskan dengan metode *global empirical Bayes* (Moraga, 2019):

$$\hat{r}_i^{EB} = C_i \frac{O_i}{n_i} + (1 - C_i)r, \quad C_i = \frac{s^2}{s^2 + \frac{r}{n_i}} \quad (2)$$

s^2 merupakan varians angka kasus antardesa yang diboboti jumlah penduduk. Desa kemudian dikelompokkan ke dalam kategori risiko rendah (*smoothed* SIR < 0,8), sedang (0,8–1,2), dan tinggi (> 1,2). Ada tidaknya pengelompokan spasial diuji dengan *global Moran's I* menggunakan matriks pembobot empat tetangga terdekat yang distandardisasi baris, dan nilai signifikansinya diperoleh melalui 9.999 kali permutasi acak.

Pemodelan Prediktif dan Evaluasi

Variabel yang diprediksi adalah jumlah kasus baru per desa untuk setiap kelompok penyakit pada bulan berikutnya. Variabel prediktor meliputi jumlah kasus 1, 2, dan 3 bulan sebelumnya, rata-rata kasus tiga bulan sebelumnya, komponen musiman sinus-kosinus, indeks waktu, kelompok penyakit, kategori wilayah, logaritma jumlah dan kepadatan penduduk, serta persentase nelayan, lansia, dan penduduk berpendidikan rendah. Model yang dibandingkan adalah *random forest*, *histogram-based gradient boosting* dengan fungsi kerugian Poisson, dan regresi Poisson. Ketiganya dibandingkan dengan model pembandingan sederhana, yaitu *naive model* (kasus bulan sebelumnya), *seasonal naive* (kasus pada bulan yang sama tahun sebelumnya), dan *moving average* tiga bulan. Data April 2024 sampai Desember 2025 digunakan untuk melatih model, sedangkan data Januari sampai Juni 2026 digunakan untuk pengujian. Kinerja model dinilai dengan *mean absolute error* (MAE), *root mean square error* (RMSE), dan *mean absolute percentage error* (MAPE) yang dihitung pada total bulanan setiap kelompok penyakit. Konsistensi hasil diperiksa dengan *rolling-origin validation* pada 12 bulan terakhir, perbedaan *absolute error* antarmodel diuji dengan *Wilcoxon signed-rank test*, dan kontribusi prediktor dinilai dengan *permutation importance* (30 ulangan). Analisis dilakukan dengan Python 3 (pandas, SciPy, scikit-learn, dan matplotlib).

HASIL DAN PEMBAHASAN

Karakteristik dan Kualitas Data

Data yang dianalisis berjumlah 42.100 baris dengan total 176.561 kunjungan dan 972 kode ICD-10 yang berbeda. Dari jumlah tersebut, 149.042 kunjungan (84,4%) merupakan kasus baru dan 27.519 kunjungan (15,6%) merupakan kasus lama. Data tersedia lengkap untuk seluruh bulan dan desa selama 30 bulan pengamatan. Angka kunjungan per 1.000 penduduk per tahun di desa daratan (841) dan kepulauan (834) hampir sama (Tabel 1), sehingga perbedaan pola penyakit tidak semata-mata disebabkan oleh perbedaan pemanfaatan layanan.

Tabel 1. Karakteristik data kunjungan empat puskesmas, Januari 2024–Juni 2026

Puskesmas	Wilayah	Desa	Penduduk 2024	Kunjungan	Kasus baru	Kode ICD-10
Bungoro	Daratan	5	31.055	56.302	53.389	721
Minasatene	Daratan	6	31.825	75.869	55.204	661
Sabutung	Kepulauan	7	14.519	32.322	30.043	426
Liukang Tupabbiring	Kepulauan	3	6.765	12.068	10.406	326
Total		21	84.164	176.561	149.042	972

Pemeriksaan kualitas data menemukan beberapa masalah. Kode gejala (R) ditemukan pada 14,3% kunjungan, terutama batuk (R05; 8.473 kunjungan), demam (R50; 6.856), dan sakit kepala (R51; 2.991), sedangkan kode Z seperti pemeriksaan kehamilan dan pemeriksaan umum mencakup 11,5% kunjungan. Beberapa label diagnosis juga tidak sesuai dengan kodenya, misalnya kode M53 yang diberi label "arthritis reumatoid" (seharusnya M05–M06) dan label kode M10 yang terpotong. Proporsi kasus lama sangat bervariasi antarpuskesmas, mulai dari 5,2% di Puskesmas Bungoro hingga 27,2% di Puskesmas Minasatene. Perbedaan ini lebih mungkin disebabkan oleh cara pengisian status kasus yang tidak seragam. Hasil ini sejalan dengan temuan ketidaktepatan pengodean di puskesmas lain (Ornela et al., 2025) dan menunjukkan bahwa tahap pemahaman data dalam CRISP-DM memegang peran penting dalam upaya optimalisasi data rutin puskesmas (Pohan, 2025).

Pola Penyakit Daratan dan Kepulauan

Sepuluh diagnosis kasus baru terbanyak didominasi oleh hipertensi, gejala gangguan pernapasan, dan gangguan pencernaan (Tabel 2). Urutannya cukup berbeda antara daratan dan kepulauan. Hipertensi menempati urutan pertama di kepulauan,

sedangkan di daratan berada di urutan ketujuh. Sebaliknya, ISPA tidak spesifik (J06) berada di urutan pertama di daratan, tetapi hanya menempati urutan ke-29 di kepulauan, sementara batuk (R05) berada di urutan kedua di kepulauan. Kondisi ini menunjukkan bahwa keluhan pernapasan di puskesmas kepulauan lebih sering dicatat dengan kode gejala dibandingkan kode diagnosis.

Tabel 2. Sepuluh diagnosis kasus baru terbanyak dan peringkatnya menurut wilayah

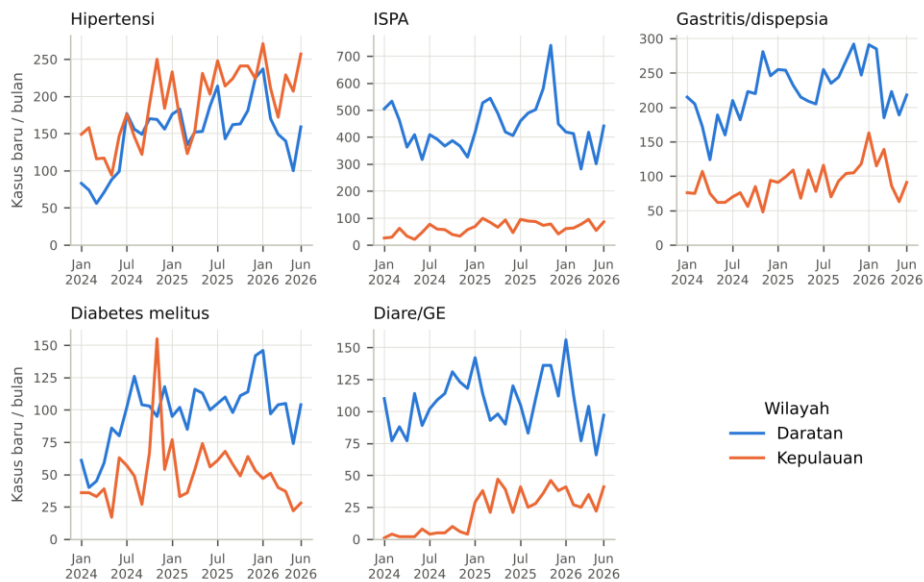
Kode	Diagnosis	Total	%	Daratan (peringkat)	Kepulauan (peringkat)
I10	Hipertensi esensial	9.776	6,6	4.241 (7)	5.535 (1)
R05	Batuk	8.171	5,5	4.558 (6)	3.613 (2)
K30	Dispepsia	7.265	4,9	6.235 (2)	1.030 (10)
J00	Nasofaringitis akut (common cold)	7.040	4,7	5.571 (3)	1.469 (8)
J06	ISPA multipel/tidak spesifik	6.982	4,7	6.733 (1)	249 (29)
R50	Demam tidak diketahui sebabnya	6.788	4,6	4.764 (5)	2.024 (4)
K04	Penyakit pulpa dan periapikal	5.877	3,9	5.007 (4)	870 (12)
M79	Gangguan jaringan lunak lainnya	4.331	2,9	1.665 (15)	2.666 (3)
Z34	Pengawasan kehamilan normal	4.247	2,8	2.288 (11)	1.959 (5)
Z00	Pemeriksaan umum tanpa keluhan	3.561	2,4	3.472 (8)	89 (56)

Kelima kelompok penyakit prioritas mencakup 28,8% dari seluruh kasus baru. Berdasarkan Tabel 3, angka hipertensi di desa kepulauan mencapai 107,1 per 1.000 penduduk per tahun, jauh lebih tinggi dibandingkan desa daratan yang sebesar 28,5 per 1.000 (RR 3,76; IK95% 3,62-3,91). Angka diabetes melitus (RR 1,55) serta gastritis dan dispepsia (RR 1,19) juga lebih tinggi di kepulauan. Kondisi sebaliknya terlihat pada ISPA (RR 0,43) dan diare (RR 0,60) yang lebih rendah di kepulauan. Hasil ini tetap sama ketika penyebut diganti dengan data penduduk Dukcapil 2025 maupun Kementerian Dalam Negeri 2020 (RR hipertensi 3,75 pada keduanya).

Tabel 3. Kasus baru, angka per 1.000 penduduk per tahun, rasio angka, dan tren tahunan lima kelompok penyakit prioritas

Kelompok penyakit	Kepulauan, kasus (angka)	Daratan, kasus (angka)	RR kepulauan/daratan (IK95%)	Perubahan per tahun, % (IK95%)
Hipertensi	5.701 (107,1)	4.477 (28,5)	3,76 (3,62–3,91)	28,6 (18,2 s.d. 39,9)
ISPA	1.896 (35,6)	13.133 (83,5)	0,43 (0,41–0,45)	5,1 (-4,9 s.d. 16,2)
Gastritis dan dispepsia	2.703 (50,8)	6.729 (42,8)	1,19 (1,13–1,24)	14,9 (9,1 s.d. 21,0)
Diabetes melitus	1.541 (29,0)	2.940 (18,7)	1,55 (1,46–1,65)	13,0 (1,2 s.d. 26,1)
Diare dan gastroenteritis	653 (12,3)	3.204 (20,4)	0,60 (0,55–0,65)	17,4 (7,6 s.d. 28,2)

Hipertensi menunjukkan tren peningkatan yang konsisten di kedua wilayah dengan perubahan sekitar 28,6% per tahun (IK95% 18,2–39,9; uji Mann-Kendall $p < 0,001$). Gastritis dan dispepsia meningkat 14,9% per tahun, sedangkan ISPA secara keseluruhan cenderung stabil (Gambar 2). Pada kasus diare di kepulauan, terjadi lonjakan dari rata-rata kurang dari 10 kasus per bulan pada tahun 2024 menjadi sekitar 20–47 kasus per bulan sejak Januari 2025. Kenaikan yang terjadi secara tiba-tiba pada satu titik waktu ini lebih mungkin disebabkan oleh perubahan cara pencatatan atau pengodean daripada peningkatan kejadian penyakit. Puncak musiman umumnya terjadi pada Oktober hingga Desember.



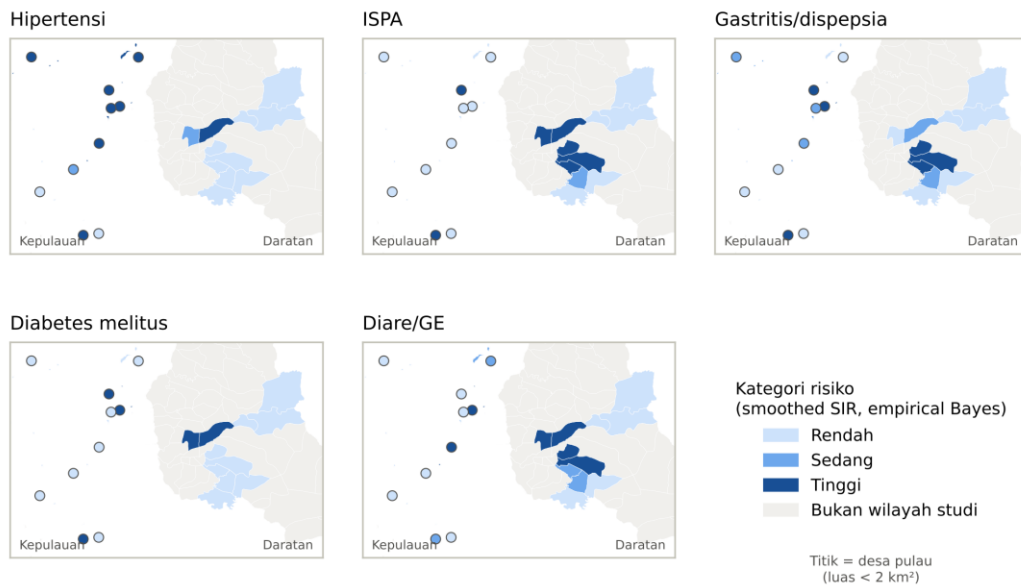
Gambar 2. Tren bulanan kasus baru lima kelompok penyakit prioritas menurut wilayah, Januari 2024–Juni 2026

Tingginya angka hipertensi di wilayah kepulauan kemungkinan dipengaruhi oleh dua faktor yang saling berkaitan. Faktor pertama adalah risiko penyakit, karena pola makan dan pekerjaan masyarakat pesisir dan nelayan dapat meningkatkan risiko hipertensi. Penelitian pada nelayan di pesisir Puger menemukan hubungan antara asupan kolesterol dan kafein dengan kejadian hipertensi (Munawaroh & Astuti, 2024), sementara di Karimunjawa sebanyak 81% nelayan yang diperiksa mengalami hipertensi (Ratih et al., 2020). Faktor kedua berkaitan dengan penemuan kasus. Layanan bergerak Perahu Sehat Pulau Bahagia yang dijalankan Puskesmas Sabutung (Budiman & Aruan, 2021) dan kegiatan skrining keliling penyakit tidak menular berpeluang menemukan kasus baru yang sebelumnya belum terdiagnosis (Firmansyah et al., 2021). Kedua faktor ini belum dapat dipisahkan dengan data agregat, tetapi keduanya menunjukkan perlunya tata laksana hipertensi yang berkelanjutan di desa pulau, termasuk ketersediaan obat dan pemantauan tekanan darah rutin.

Rendahnya angka ISPA di kepulauan perlu dimaknai secara hati-hati. Hambatan akses dapat membuat masyarakat tidak datang ke puskesmas untuk keluhan ringan (Nurfadillah et al., 2025; Siregar et al., 2025). Di samping itu, kode batuk (R05) jauh lebih sering digunakan di puskesmas kepulauan, yang menunjukkan adanya kebiasaan mengode berdasarkan gejala. Jika kode R05 ikut diperhitungkan, perbedaan angka ISPA antara daratan dan kepulauan kemungkinan akan mengecil. Karena itu, *dashboard* perlu menampilkan indikator kualitas data seperti proporsi kode gejala agar perbedaan pengodean tidak disalahartikan sebagai perbedaan kejadian penyakit.

Pemetaan Risiko Tingkat Desa

Hasil pemetaan menunjukkan adanya pengelompokan spasial yang bermakna pada hipertensi (*Moran's I* 0,487; $p < 0,001$), ISPA (0,249; $p = 0,025$), dan diare (0,226; $p = 0,033$), tetapi tidak pada gastritis dan dispepsia (0,135; $p = 0,074$) maupun diabetes melitus ($-0,031$; $p = 0,368$). Pada hipertensi, tujuh dari delapan desa dengan kategori risiko tinggi berada di wilayah kepulauan, dengan nilai *smoothed SIR* tertinggi di Mattiro Kanja (6,70), Mattiro Baji (4,78), dan Mattiro Sompe (2,49). Sebaliknya, desa dengan risiko tinggi ISPA dan diare lebih banyak berada di daratan, yaitu Minasatene (*SIR* 2,28 dan 1,99), Samalewa (1,63 dan 1,72), dan Biraeng (1,63 untuk ISPA). Mattiro Kanja masuk kategori risiko tinggi untuk tiga kelompok penyakit sekaligus, yaitu hipertensi, diabetes melitus (4,16), serta gastritis dan dispepsia (3,18) (Gambar 3).



Gambar 3. Kategori risiko lima kelompok penyakit prioritas menurut desa berdasarkan SIR yang dihaluskan dengan *empirical Bayes*

Temuan pengelompokan spasial ini sejalan dengan penelitian berbasis wilayah kerja puskesmas yang menunjukkan bahwa kasus penyakit cenderung terkonsentrasi di lokasi tertentu sehingga dapat digunakan untuk menentukan wilayah prioritas intervensi (Syamsir et al., 2020; Hidayati, 2023). Namun, penelitian ini hanya mencakup 21 desa yang terbagi dalam dua kelompok wilayah yang terpisah, sehingga nilai *Moran's I* lebih banyak menggambarkan perbedaan antara daratan dan kepulauan daripada kluster lokal di dalam masing-masing wilayah. Karena itu, peta risiko yang dihasilkan lebih tepat digunakan untuk menentukan desa prioritas daripada untuk menyimpulkan adanya proses penularan antarwilayah. *Empirical Bayes smoothing* diperlukan karena desa pulau berpenduduk sekitar 1.000–2.000 jiwa menghasilkan angka mentah yang sangat berfluktuasi (Moraga, 2019).

Kinerja Model Prediksi

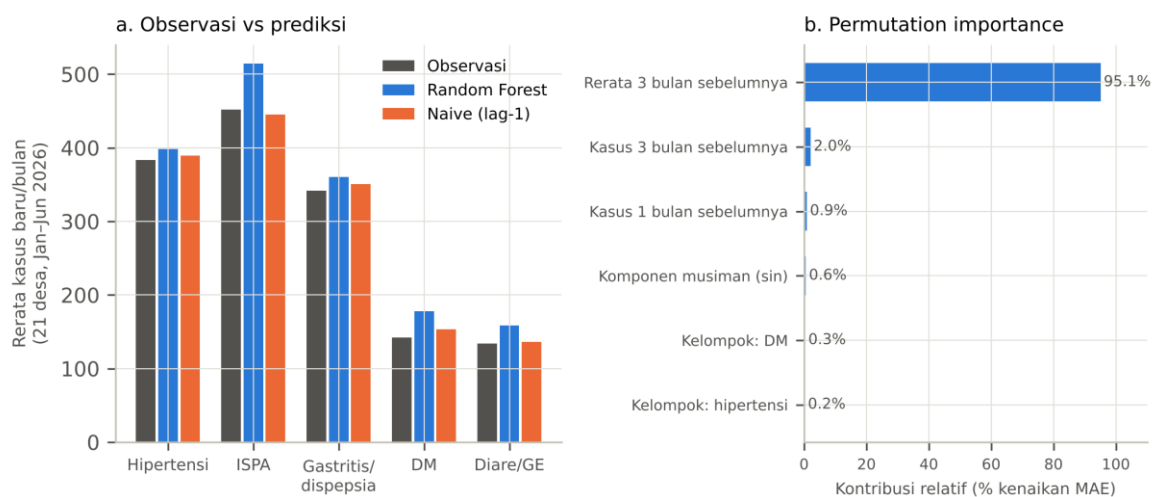
Tabel 4 menunjukkan bahwa *random forest* memiliki RMSE paling rendah (7,99) dengan MAE sebesar 4,88 kasus per desa per bulan. Hasil ini hanya sedikit lebih baik dibandingkan *naive model* (MAE 5,10) dan hampir sama dengan *moving average* tiga bulan (MAE 4,84). Menurut *Wilcoxon signed-rank test*, perbedaan *absolute error* antara *random forest* dan *naive model* tidak bermakna ($p = 0,111$). Pada *rolling-origin validation* selama 12 bulan, *random forest* menurunkan rata-rata MAE dari 5,01 menjadi 4,67 kasus atau sekitar 6,8%. Penurunan galat terlihat pada hipertensi (MAE 6,06 dibandingkan 6,90), gastritis dan dispepsia (4,80 dibandingkan 5,59), serta diare (2,88 dibandingkan 2,97), tetapi tidak

terlihat pada ISPA dan diabetes melitus. Regresi Poisson menghasilkan galat terbesar karena hasilnya tidak stabil pada beberapa desa yang mengalami lonjakan kasus.

Tabel 4. Kinerja model prediksi kasus baru bulanan pada periode uji Januari–Juni 2026

Model	MAE	RMSE	MAPE agregat (%)
Naive (kasus bulan lalu)	5,10	8,36	23,4
Seasonal naive (bulan sama tahun lalu)	5,98	10,34	20,3
Moving average 3 bulan	4,84	8,01	23,0
Regresi Poisson	8,26	24,71	26,2
Gradient boosting (Poisson)	4,93	8,26	22,1
Random forest	4,88	7,99	23,5

Berdasarkan hasil *permutation importance*, rata-rata kasus tiga bulan sebelumnya memberikan kontribusi sebesar 95,1% terhadap total kenaikan galat ketika nilai prediktor diacak, jauh lebih besar dibandingkan variabel sosiodemografi dan variabel musiman (Gambar 4b). Pada semester pertama 2026, *random forest* cenderung memprediksi lebih tinggi daripada kasus teramati karena jumlah kasus menurun dibandingkan akhir 2025 (Gambar 4a).



Gambar 4. Perbandingan kasus teramati dan prediksi pada periode uji (a) serta kontribusi relatif prediktor model *random forest* (b)

Model yang lebih kompleks tidak selalu menghasilkan kinerja prediksi yang lebih baik, terutama ketika diterapkan pada data deret waktu kesehatan dengan jumlah observasi yang terbatas. Dalam studi di Bandung, kinerja model *machine learning* dan *neural network* menurun tajam pada data uji, sedangkan model Bayesian spasio-temporal

memberikan hasil yang paling stabil (Jaya et al., 2025). Ajulo et al. (2025) juga menekankan pentingnya membandingkan model dengan pembanding sederhana serta melaporkan proses validasi secara transparan. Monteiro et al. (2025) melaporkan bahwa penggabungan *spatial clustering* dengan model prediksi dapat meningkatkan akurasi pada wilayah yang heterogen. Akan tetapi, dengan data 30 bulan, informasi yang dimanfaatkan model sebagian besar berasal dari kasus bulan-bulan terakhir. Bagi puskesmas, hal ini memudahkan karena *moving average* tiga bulan dapat dihitung langsung oleh petugas. *Random forest* tetap dapat digunakan sebagai pelengkap, khususnya untuk hipertensi dan gastritis yang menunjukkan perbaikan akurasi.

Implikasi bagi Perencanaan Program dan Dashboard

Program pengendalian penyakit tidak menular di Puskesmas Sabutung dan Puskesmas Liukang Tupabbiring perlu memprioritaskan desa dengan SIR hipertensi tinggi, terutama Mattiro Kanja, Mattiro Baji, dan Mattiro Sompe, misalnya dengan memasukkan pemantauan tekanan darah dan distribusi obat ke dalam jadwal layanan bergerak. Program pencegahan ISPA dan diare sebaiknya difokuskan pada desa daratan yang padat penduduk, seperti Minasatene, Biraeng, dan Samalewa. Selain itu, pelatihan pengodean diagnosis dan pengisian status kasus perlu menjadi bagian dari upaya optimalisasi data, karena kualitas pengodean akan menentukan ketepatan peta risiko. Komponen tersebut telah dituangkan ke dalam *dashboard* berisi peta risiko desa, tren bulanan, prediksi bulan berikutnya, dan indikator kualitas data, sejalan dengan pengalaman penerapan *dashboard* di beberapa dinas kesehatan (Hizriansyah, 2023; Sari et al., 2022; Putri & Ramani, 2024) dan pemanfaatan visualisasi data untuk mengenali pola kunjungan pasien (Akbar et al., 2025; Rindiyan et al., 2025).

Penelitian ini memiliki beberapa keterbatasan. Data penduduk yang digunakan sebagai penyebut hanya berasal dari satu titik waktu (2024). Pemilihan empat puskesmas secara *purposive* belum dapat mewakili seluruh wilayah kabupaten, dan periode pengamatan selama 30 bulan membatasi kemampuan model dalam menangkap pola musiman. Faktor lingkungan, seperti iklim dan waktu tempuh ke fasilitas kesehatan, juga belum dimasukkan ke dalam model.

SIMPULAN DAN SARAN

Data rutin dari empat puskesmas di Kabupaten Pangkajene dan Kepulauan dapat diolah menjadi informasi spasio-temporal yang bermanfaat bagi perencanaan program, dengan ketentuan bahwa data tersebut melalui proses pemahaman dan persiapan yang

sistematis. Pola penyakit di wilayah daratan dan kepulauan menunjukkan perbedaan yang jelas. Hipertensi, diabetes melitus, serta gastritis dan dispepsia lebih tinggi dan cenderung mengelompok di desa kepulauan, sedangkan ISPA dan diare lebih tinggi di desa daratan yang padat penduduk. Kasus hipertensi juga terus meningkat selama periode penelitian. Model *random forest* hanya memberikan sedikit perbaikan akurasi dibandingkan *naive model* dan hasilnya setara dengan *moving average* tiga bulan. Model prediksi sederhana sudah cukup untuk digunakan dalam *dashboard* perencanaan puskesmas.

Dinas Kesehatan disarankan untuk memanfaatkan peta risiko desa dalam menentukan alokasi program, memberikan prioritas pada pengelolaan hipertensi di desa-desa pulau, serta melaksanakan pelatihan pengodean ICD-10 bagi petugas puskesmas. Penelitian selanjutnya perlu mencakup seluruh puskesmas di Kabupaten Pangkep, menggunakan data penduduk yang lebih mutakhir, memasukkan faktor lingkungan dan aksesibilitas, serta menerapkan model spasio-temporal hierarkis agar pengaruh kejadian penyakit dapat dipisahkan dari pengaruh penemuan kasus.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Kementerian Pendidikan Tinggi, Sains, dan Teknologi atas pendanaan penelitian melalui skema Penelitian Dosen Pemula tahun 2026 dengan nomor kontrak induk 288/C3/DT.05.00/PL-BARU/2026 dan kontrak turunan 407/LL9/PPIP-PL.MONO/2026 serta 131/A-VI/UNIPRIMA/I/V/2026. Ucapan terima kasih juga disampaikan kepada Lembaga Penelitian, Pengabdian Masyarakat dan Penjaminan Mutu Universitas Puangrimaggalatung, Dinas Kesehatan Kabupaten Pangkajene dan Kepulauan serta Puskesmas Bungoro, Puskesmas Minasatene, Puskesmas Sabutung, dan Puskesmas Liukang Tupabbiring atas dukungan dalam penyediaan data.

DAFTAR PUSTAKA

- Ajulo, H. K., Alele, F. O., Emeto, T. I., & Adegboye, O. A. (2025). Spatial And Spatiotemporal Machine Learning Models for COVID-19 Dynamics: A Review of Methodology and Reporting Practices. *Epidemiologic Reviews*, 47(1), mxaf017. <https://doi.org/10.1093/epirev/mxaf017>
- Akbar, F., Rafiqah, & Nirad, D. W. S. (2025). Implementasi Dashboard Pasien Rawat Jalan Rumah Sakit Umum XYZ dengan Fitur *Clustering*, *Forecasting* dan *Outlier Detection*. *Jurnal Nasional Teknologi dan Sistem Informasi (TEKNOSI)*, 11(2), 170–177. <https://doi.org/10.25077/TEKNOSI.v11i2.2025.170-177>
- Andayani, S., Marwoto, B. S. H., Yusri, T. A. S., Setyaningsih, F. A., & Lathifah, F. A. (2025). Transformasi Data Kesehatan: Pelatihan Penyusunan Dashboard Interaktif di Dinas Kesehatan Gunung Kidul. *Abdi: Jurnal Pengabdian dan Pemberdayaan Masyarakat*, 7(3), 804–813. <https://doi.org/10.24036/abdi.v7i3.1252>

- Badan Pusat Statistik Kabupaten Pangkajene dan Kepulauan. (2025). *Kecamatan dalam Angka 2025: Bungoro, Minasatene, Liukang Tupabbiring, dan Liukang Tupabbiring Utara*. Pangkajene: BPS Kabupaten Pangkajene dan Kepulauan.
- Budiman, V., & Aruan, M. (2021). *Perahu Sehat Pulau Bahagia (PSPB): Layanan Kesehatan Dasar Mobile untuk Menjangkau Masyarakat Kepulauan*. Seri Dukungan KOMPAK untuk Kebijakan Nasional. Jakarta: KOMPAK.
- Dinas Kependudukan dan Pencatatan Sipil Kabupaten Pangkajene dan Kepulauan. (2026). *Jumlah Keluarga dan Penduduk berdasarkan Data Konsolidasi Bersih Tahun 2025 Semester 2*. Pangkajene: Disdukcapil Kabupaten Pangkajene dan Kepulauan.
- Dinas Kesehatan Kabupaten Pangkajene dan Kepulauan. (2025). *Rencana Strategis Dinas Kesehatan Kabupaten Pangkajene dan Kepulauan Tahun 2025–2029*. Pangkajene: Dinas Kesehatan Kabupaten Pangkajene dan Kepulauan.
- Firmansyah, Y., Ginting, D. N., Su, E., Sylvana, Y., Chau, W., & Setyati, P. N. (2021). Pentingnya Posbindu Keliling dalam Mendeteksi Penyakit Tidak Menular khususnya Hipertensi di RW 05, Kelurahan Kedaung Kaliangke. *Jurnal Muara Sains, Teknologi, Kedokteran, dan Ilmu Kesehatan*, 5(1), 9–18. <https://doi.org/10.24912/jmstkik.v5i1.6344>
- Hidayati, N. (2023). Analisis Spasial Kejadian Demam Berdarah Dengue (DBD) di Kabupaten Bantul Tahun 2022. *Jurnal Kesehatan Masyarakat Indonesia*, 18(4), 27–33. <https://doi.org/10.26714/jkmi.18.4.2023.27-33>
- Hizriansyah, H. (2023). Perancangan Model Dashboard untuk Pelaporan dan Visualisasi Data Kesehatan sebagai Sistem Monitoring di Dinas Kesehatan Gunungkidul. *Journal of Information Systems for Public Health*, 8(1), 1–9. <https://doi.org/10.22146/jisph.72268>
- Jaya, I. G. N. M., Andriyana, Y., Tantular, B., Pangastuti, S. S., & Kristiani, F. (2025). Spatiotemporal Dengue Forecasting for Sustainable Public Health in Bandung, Indonesia: A Comparative Study of Classical, Machine Learning, and Bayesian Models. *Sustainability*, 17(15), 6777. <https://doi.org/10.3390/su17156777>
- Kementerian Dalam Negeri. (2020). *Data Batas Administrasi Desa dan Agregat Kependudukan per Desa/Kelurahan Semester I Tahun 2020*. Jakarta: Direktorat Jenderal Kependudukan dan Pencatatan Sipil.
- Monteiro, K. H. de C., Rocha, É. da S., Morais, L. A., Santos, E. G., Neto, S. R. da S., Sampaio, V., & Endo, P. T. (2025). Integrating Machine Learning and Spatial Clustering for Malaria Case Prediction in Brazil's Legal Amazon. *BMC Infectious Diseases*, 25, 802. <https://doi.org/10.1186/s12879-025-11193-x>
- Moraga, P. (2019). *Geospatial health data: Modeling and Visualization with R-INLA and Shiny*. Boca Raton: Chapman & Hall/CRC. <https://doi.org/10.1201/9780429341823>
- Munawaroh, N. W., & Astuti, N. F. W. (2024). Pola Konsumsi Natrium, Kolesterol, dan Kafein terhadap Kejadian Hipertensi pada Nelayan di Pesisir Puger Jember. *Journal of Nutrition College*, 13(2), 177–187. <https://doi.org/10.14710/jnc.v13i2.39430>
- Nurfadillah, Hikmah B, N., & Andayanie, E. (2025). Faktor yang Berhubungan terhadap Akses Pelayanan Kesehatan di Puskesmas Seko Desa Padang Raya. *Window of Public Health Journal*, 6(2), 330–341.
- Ornela, O., Putra, D. M., Novita, D., Wisandra, A., & Leonard, D. (2025). Tinjauan Ketepatan Kode Diagnosis ICD-10 Diabetes Melitus pada Rekam Medis Rawat Jalan di Puskesmas Taram Kabupaten Lima Puluh Kota. *Jurnal Rekam Medis dan Informasi Kesehatan (J-REMIKES)*, 4(1), 29–34.

- Pohan, H. I. (2025). *Constructing A Dataset for Infectious Disease Prediction And Spatial Cluster Analysis*. *J-Icon: Jurnal Komputer dan Informatika*, 13(2), 60–67. <https://doi.org/10.35508/jicon.v13i2.23729>
- Putri, E. F. K., & Ramani, A. (2024). Perancangan Dashboard Visualisasi Data Kesehatan Ibu dan Anak di Dinas Kesehatan Kabupaten Jember. *e-Journal Pustaka Kesehatan*, 12(1).
- Rahmi, A., Lastri, S., & Hasnur, H. (2024). PIECES (*Performance, Information, Economic, Control, Efficiency, Service*) dengan Pemanfaatan Sistem Informasi Manajemen Puskesmas (SIMPUS). *Jambura Journal of Health Science and Research*, 6(2), 146–154.
- Ratih, P., Bayu, Y. S. N., & Haikal. (2020). Gambaran Kejadian Hipertensi pada Nelayan di Desa Kemojan, Karimun Jawa, Jepara. *Visikes: Jurnal Kesehatan Masyarakat*, 19(1), 287–295.
- Rindiyan, F. D., Fauzi, H., Nadi, D. T., Hakim, A. O., & Aeni, Y. N. (2025). Visualisasi Spasial Kunjungan Pasien Rumah Sakit Menggunakan Sistem Informasi Geografis. *J-REMI: Jurnal Rekam Medik dan Informasi Kesehatan*, 6(3), 168–184. <https://doi.org/10.25047/j-remi.v6i3.5866>
- Sari, R. Y., Supriati, T., Pratama, R. A., & Sutriana, V. N. (2022). Model Dashboard Informasi Pendukung Program Kesga: Studi Implementasi di Dinkes Kabupaten Kulon Progo. *Journal of Information Systems for Public Health*, 7(1), 1–10.
- Schröer, C., Kruse, F., & Gómez, J. M. (2021). *A Systematic Literature Review on Applying CRISP-DM Process Model*. *Procedia Computer Science*, 181, 526–534. <https://doi.org/10.1016/j.procs.2021.01.199>
- Siregar, P. P., Bhuwana, S. C., & Toniara, S. (2025). Faktor yang Mempengaruhi Masyarakat Terhadap Akses Menuju Puskesmas Lubuk Pakam. *Jurnal Pandu Husada*, 6(3), 21–29.
- Syamsir, S., Daramusseng, A., & Rudiman, R. (2020). Autokorelasi Spasial Demam Berdarah Dengue di Kecamatan Samarinda Utara, Kota Samarinda. *Jurnal Kesehatan Lingkungan Indonesia*, 19(2), 119–126. <https://doi.org/10.14710/jkli.19.2.119-126>
- Thenu, V. J., Sedyono, E., & Purnami, C. T. (2016). Evaluasi Sistem Informasi Manajemen Puskesmas Guna Mendukung Penerapan SIKDA Generik menggunakan Metode HOT-fit di Kabupaten Purworejo. *Jurnal Manajemen Kesehatan Indonesia*, 4(2), 129–138.
- Yani, B. K. D., & Lazuardi, L. (2018). Fungsi Esensial Sistem Informasi Manajemen Puskesmas (SIMPUS) di Kota Yogyakarta. *Journal of Information Systems for Public Health*, 3(3), 78–84.