1. Title Page

Title	Use of metamizole and the risk of acute kidney injury
Research question & Objectives	To determine if the use of metamizole in postoperative settings causes an increased risk of acute kidney injury in inpatients compared to opioids and NSAIDs.
Protocol version	V1.3
Last update date	24/09/2025
Contributors	**Primary investigator contact information:** Daniel Fernández-Llaneza **Contributor names:** Izak Yasrebi-de Kom, Daniala L. Weir, Ameen Abu-Hanna, Giovanni Cinà, Joanna E. Klopotowska
Study registration	**Site:** HMA/EMA RWD Catalogues **Identifier:** EUPAS1000000740
Sponsor	**Organization:** Amsterdam UMC/Dutch Research Council (Nederlandse Organisatie voor Wetenschappelijk Onderzoek, NWO) **Contact:** Joanna E. Klopotowska
Conflict of interest	n/a

Table of contents

2. Abstract

Metamizole (dipyrone) was first introduced commercially in 1922 an effective non-opioid analgesic and antipyretic, indicated for moderate to severe pain, particularly in postoperative settings. Although concerns about serious adverse effects like agranulocytosis have persisted, its safety profile remains debated. It is recommended by the Dutch Association of Anaesthesiologists (Nederlandse Vereniging voor Anesthesiologie, NVA), as a suitable alternative for patients who cannot take non-steroidal anti-inflammatory drugs (NSAIDs). However, despite its widespread use, its acute kidney injury (AKI) potential has received limited attention. This study will evaluate the AKI potential of metamizole administered in postoperative settings in a major academic hospital in the Netherlands to provide evidence to inform clinical practice.

We will use electronic health records (EHR) of patients who were hospitalised at the Amsterdam University Medical Centre (UMC), which includes longitudinal date-stamped outpatient and inpatient information on social demographics, laboratory measurements, healthcare utilisation, and medical diagnoses. We will carry out a cohort study with a new user active comparator design, where metamizole new-users will be compared against opioid new-users and NSAIDs new-users within the hospital premises.

3. Amendments and updates

Version date	Version number	Section of protocol	Amendment or update	Reason
13 May 2025	V1	First draft	n/a	n/a
01 Jul 2025	V1.1	Second draft	Update after comments from the rest of the team	n/a
24 Sep 2025	V1.2	Third draft	Last review prior to starting the formal analysis	n/a

4. Milestones

Table 1 Milestones

Milestone	Date
Feasibility counts	8 May 2025
Draft 1 of protocol complete	24 September 2025
Registration of protocol	
Study progress report 1	
Study progress report 2	
Final report of study results	

5. Rationale and background

What is known about the condition: Acute kidney injury (AKI) is defined as a sudden decline in kidney function identified via increased serum creatinine levels or reduced urine output. Over the past decades, AKI has been linked with poor health outcomes even at milder stages and occur approximately in 20% of hospitalised adults. It is estimated that 19-26% of AKI events in hospitalised patients are due to drugs.

What is known about the exposure of interest: Metamizole (dipyrone) is used as a pain reliever in postoperative settings and, in the Netherlands, it is recommended for use in patients with a contraindication to NSAIDs. Its international status with regard to regulatory approval remains controversial due to the risk of agranulocytosis, a potentially fatal adverse drug event. It is suspected that metamizole may also be causative of AKI, but current evidence is still scarce and with methodological limitations. Evidence from current randomised clinical trials reported similar kidney toxicity profiles to naproxen, diclofenac or paracetamol.

Gaps in knowledge: the causal relationship between metamizole use and acute kidney injury remains controversial.

What is the expected contribution of this study? To provide further evidence on the acute kidney injury potential of metamizole.

6. Research question and objectives

Table 2 Primary and secondary research questions and objective

A. Primary research question and objective

Objective:	To estimate the *total effect* of metamizole on the risk of acute kidney injury compared with (1) opioids and (2) NSAIDs, allowing all intercurrent events and censoring to occur as observed in clinical practice.
Hypothesis:	Metamizole use increases the risk of acute kidney injury compared to opioids and may have a similar or lower risk compared to NSAIDs, when accounting for the natural course of treatment and patient outcomes in routine care.
Population *(mention key inclusion-exclusion criteria):*	First ever inpatient administration of metamizole users in a postoperative setting (i.e., up to 7 days after surgery). Patients who were prescribed metamizole up to 14 days prior to surgery were excluded, as they were likely to be prevalent users. Patients on multimodal analgesia (i.e., receiving more than one pain relief medication) within 1 hour of surgery were excluded, because it was considered that these medications were intended to be given together. Patients with acute kidney injury or acute kidney disease episodes prior up to one week prior to surgery were excluded.
Exposure:	Metamizole (dipyrone)
Comparator:	Non-steroidal anti-inflammatory agents
Outcome:	Acute kidney injury
Time *(when follow up begins and ends):*	Follow-up for the cohort began one hour after initiation of therapy with metamizole (or comparator) until the first-ever acute kidney injury during hospital stay, identification of pregnancy, receipt of kidney replacement therapy (i.e., chronic dialysis, kidney transplant), treatment strategy switch/add-on, in-hospital death, hospital discharge, 14 days after initiation, or end of the study period. All incident cases of acute kidney injury occurring during follow-up were identified. Exposure to metamizole was evaluated any time from surgery until 7 days after surgery.
Setting:	Inpatient care.
Main measure of effect:	risk difference (primary estimand), risk ratio (supplementary estimand)

a. Estimands framework

Estimand Attribute	Estimand 1 (primary estimand)	Estimand 2 (supplementary estimand)	Estimand 3 (supplementary estimand)	Estimand 4 (supplementary estimand)
Population	Adults hospitalised in Amsterdam UMC for ≥24 hours between 2019 and 2024 in a postoperative setting (i.e., 7 days from surgery).			
Treatment Conditions	Intervention group: first-ever systemic administration of metamizole. Control group: first-ever systemic administration of opioids (i.e., morphine, buprenorphine, oxycodone, piritramide). In-class switches or dose adjustments during follow-up are considered part of the same initial treatment strategy and do not constitute treatment changes.		Intervention group: first-ever systemic administration of metamizole. Control group: first-ever systemic administration of non-steroidal anti-inflammatory drugs. In-class switches or dose adjustments during follow-up are considered part of the same initial treatment strategy and do not constitute treatment changes.	
Endpoint	First-ever inpatient acute kidney injury, based on KDIGO criteria (i.e., serum creatinine).			
Summary Measure	Risk difference	Risk ratio	Risk difference	Risk ratio
Intercurrent events and strategies to handle them*	Same for both treatment conditions. Intercurrent events: 1) Receipt of kidney replacement therapy (i.e., chronic dialysis or kidney transplantation): treatment policy. 2) Out-of-class treatment switch/add-on: treatment policy. 3) Documentation of pregnancy: treatment policy. 4) Administration of drug(s) with (drug–drug) interactions flagged for AKI: treatment policy. 5) In-hospital death: treatment policy. 6) Use of prohibited medications (i.e., selected drugs with AKI potential): treatment policy. 7) Prolonged analgesic rescue medication (>48 hours): treatment policy. 8) Treatment discontinuation: treatment policy.			

* Intercurrent events are post-baseline events (or post-randomisation events in randomised trials) that affect the interpretation or existence of outcome data. These events frequently affect receipt of treatment (e.g., treatment switching or treatment discontinuation) or preclude existence of the outcome (e.g., death, if it is not defined as part of the outcome).

B. Secondary research question and objective

Objective:	To estimate the *controlled direct effect* of metamizole on the risk of acute kidney injury, compared with (1) opioids and (2) non-steroidal anti-inflammatory drugs (NSAIDs), under a hypothetical scenario in which censoring occurs at random, no intercurrent events occur, regardless of treatment discontinuation.
Hypothesis:	Under conditions where no intercurrent or censoring events occur, metamizole use increases the risk of acute kidney injury compared to opioids, and has a comparable or lower risk to NSAIDs.

Population *(mention key inclusion-exclusion criteria):*	First ever inpatient administration of metamizole users in a postoperative setting (i.e., up to 7 days after surgery). Patients who were prescribed metamizole up to 14 days prior to surgery were excluded, as they were likely to be prevalent users. Patients on multimodal analgesia (i.e., receiving more than one pain relief medication) within 1 hour of surgery were excluded, because it was considered that these medications were intended to be given together. Patients with acute kidney injury or acute kidney disease episodes up to one week prior to surgery were excluded. Patients who had acute dialysis 14 days prior to exposure or on chronic dialysis were excluded, too.
Exposure:	Metamizole (dipyrone)
Comparator:	Opioids (i.e., morphine, buprenorphine, oxycodone, and piritramide), NSAIDs (i.e., aceclofenac, dexketoprofen, diclofenac, phenylbutazone, flurbiprofen, ibuprofen, indomethacin, ketoprofen, meloxicam, nabumetone, naproxen, etoricoxib, parecoxib, celecoxib, piroxicam, propyphenazone, tiaprofenic acid, high-dose acetylsalycilic acid, and diflusinal).
Outcome:	Acute kidney injury as defined by the Kidney Disease: Improving Global Outcomes (KDIGO) using serum creatinine measurements.
Time *(when follow up begins and ends):*	Follow-up for the cohort began one hour after initiation of therapy with metamizole (or comparator), this to allow a grace period in which intended multimodal therapy could be initiated (i.e., time zero). Follow-up continued until the earliest of: first-ever acute kidney injury during hospital stay, hospital discharge, 14 days after time zero, or end of study period. Patients were followed up regardless of changes in treatment status (i.e., treatment discontinuation, dose adjustments), consistent with a treatment-policy approach. Exposure to metamizole was evaluated any time from surgery until 7 days after surgery, which was considered to be postoperative period.
Setting:	Inpatient care.
Main measure of effect:	risk difference (primary estimand), risk ratio (supplementary estimand)

b. Estimands framework

Estimand Attribute	**Estimand 1 (primary estimand)**	**Estimand 2 (supplementary estimand)**	**Estimand 3 (supplementary estimand)**	**Estimand 4 (supplementary estimand)**
Population	Adults hospitalised in Amsterdam UMC for ≥24 hours between 2019 and 2024 in a postoperative setting (i.e., 7 days from surgery).			
Treatment Conditions	Intervention group: first-ever systemic administration of metamizole. Control group: first-ever systemic administration of opioids. In-		Intervention group: first-ever systemic administration of metamizole. Control group: first-ever systemic administration of non-steroidal anti-inflammatory drugs.	

Estimand Attribute	Estimand 1 (primary estimand)	Estimand 2 (supplementary estimand)	Estimand 3 (supplementary estimand)	Estimand 4 (supplementary estimand)
	class switches or dose adjustments during follow-up are considered part of the same initial treatment strategy and do not constitute treatment changes.		In-class switches or dose adjustments during follow-up are considered part of the same initial treatment strategy and do not constitute treatment changes.	
Endpoint	First-ever inpatient acute kidney injury, based on KDIGO criteria (i.e., serum creatinine).			
Summary Measure	Risk difference	Risk ratio	Risk difference	Risk ratio
Intercurrent events and strategies to handle them*	Same for both treatment conditions. Intercurrent events: 1) Receipt of kidney replacement therapy (i.e., chronic dialysis or kidney transplantation): hypothetical. 2) Out-of-class treatment switch/add-on: hypothetical. 3) Documentation of pregnancy: hypothetical. 4) Administration of drug(s) with (drug–drug) interactions flagged for AKI: hypothetical. 5) In-hospital death: hypothetical. 6) Use of prohibited medications (i.e., selected drugs with AKI potential): hypothetical. 7) Prolonged analgesic rescue medication (>48 hours): hypothetical. 8) Treatment discontinuation: treatment policy.			

* Intercurrent events are post-baseline events (or post-randomisation events in randomised trials) that affect the interpretation or existence of outcome data. These events frequently affect receipt of treatment (e.g., treatment switching or treatment discontinuation) or preclude existence of the outcome (e.g., death, if it is not defined as part of the outcome).

C. Target Trial Emulation Framework

Protocol Element		Target Trial Specification	Emulation with Observational Data
Causal Estimand	Eligibility criteria	**Inclusion criteria:** 1) Adults (≥18 years old at admission date) admitted to Amsterdam UMC for more than 24 hours; 2) Post-operative pain management setting. **Exclusion criteria:** 1) Kidney transplantation in the past year from index date, because it alters kidney physiology and the kidney function may remain unstable; 2) On chronic dialysis, because no significant remaining kidney function can sustain injury; 3) Acute dialysis episode within the past 14 days (as per ADQI 16 Working group)[1] from index date, because kidney function may be unstable or patients may still be recovering from an unresolved kidney injury; 4) Pre-exposure AKI or AKD episodes seven days prior to or on index date; 5) No metamizole or comparator use (as pertinent in each treatment strategy) 14 days prior to treatment initiation, to represent 'new user' status (washout window); 6) Pregnant women, because pregnancy-related AKI is most likely due to other reasons; 7) Patients with a treatment switch or add-on within 1 hour of the initial administration of a pain management drug (i.e., metamizole, NSAIDs, or opioids), as this likely reflects multimodal analgesia strategy rather than a true switch or add-on.	Same as target trial. Observation period between January 2019, 1 and December 2024, 31. **Inclusion criteria:** 1) Any event considered as registered in Amsterdam UMC EHR systems; 2) Post-operative pain management setting: identified via procedure codes performed in the operation room, excluding echographies. **Exclusion criteria:** 1) No kidney transplantation: defined as absence of procedure codes or ICD-10 diagnosis codes. 2) Chronic dialysis is defined as two dialysis encounters at least 90 days apart with at least two dialysis encounters per week (i.e., dialysis measurements, procedure codes or ICD-10 codes) or the equivalent total (i.e., 25 encounters in any 90-day time window) or a registration of a diagnosis code pertaining to chronic dialysis. 3) Acute dialysis episode: defined as the presence of a dialysis encounter, procedure code or ICD-10 diagnosis code 14 days prior to inpatient metamizole administration. 4) No previous AKI or AKD episode: defined as per 2012 AKI KDIGO. 5) No metamizole or comparator use (as pertinent in each treatment strategy) identified with ATC code and internal codes when the ATC code was lacking. 6) Pregnancy: identified via pregnancy status as recorded by clinician in the Pregnancy Detailed Clinical Model (DCM). A pregnancy is considered active if the recorded status falls within 294 days (approximately 42 weeks) before index date. 7) Multimodal analgesia: defined as receiving more than one pain management strategy (i.e., opioids, NSAIDs or metamizole) within 1 hour of index

Protocol Element		Target Trial Specification	Emulation with Observational Data
			administration (i.e., first-ever administration of opioids, NSAIDs or metamizole). Unlike in the target trial, given that we implement these exclusion criteria based on ICD-10, ATC, CBV, internal codes or registrations (e.g., pregnancy records), misclassification may occur. In addition, no prior AKI relies on a retrospective implementation of the KDIGO guideline on AKI, whereas in the target trial it could be done using real-time lab measurements.
	Treatment strategies	Systemic administrations are defined as any oral, parenteral or other administration routes as specified in *Supplementary Information*. (1) First-ever initiation of systemic inpatient administration of metamizole only, as-needed, according to therapeutic dosing guidelines. (2) Active comparator arms: a. Reference cohort (negative exposure): i. First-ever initiation of systemic inpatient administration of opioids (i.e., morphine, buprenorphine, oxycodone, and piritramide), according to therapeutic dosing guidelines. b. Additional cohort (positive exposure): i. First-ever initiation of systemic inpatient administration of NSAIDs (i.e., aceclofenac, dexketoprofen, diclofenac, phenylbutazone, flurbiprofen, ibuprofen, indomethacin, ketoprofen, meloxicam, nabumetone, naproxen, etoricoxib, parecoxib, celecoxib, piroxicam, propyphenazone, tiaprofenic acid, high-dose acetylsalicylic acid [i.e., ≥100 mg/day], and diflusinal), according to therapeutic dosing guidelines.	Same as target trial, identified in the Medication Administration DCM with ATC codes. In the target trial, first-ever systemic initiation is enforced and ensures a clean exposure contrast, whereas in the emulation drug exposure is reconstructed based on registrations as EHR data is re-used. Additionally, multimodal analgesia treatment patterns need to be discarded.
	Treatment Assignment (unmasked)	Randomised, non-blinded.	Non-blinded and assumed to be randomised conditional on the measured confounders. Randomisation is emulated

Protocol Element		Target Trial Specification	Emulation with Observational Data
			through inverse probability of treatment weighting (IPTW). Medications are assessed using the medication administration records and internal codes if ATC codes are missing.
	Start/end Follow-up	Follow-up starts at treatment administration initiation and is censored at *first*-ever occurrence of AKI during hospitalisation, receipt of kidney replacement therapy (i.e., chronic dialysis or kidney transplantation), 14 days after treatment initiation, out-of-class treatment strategy switch/add-on, documentation of pregnancy, in-hospital death, drug-drug interactions, prohibited medications, treatment discontinuation, hospital discharge, or administrative censoring (i.e., December 2024, 31), whichever occurs first.	Same as target trial, follow-up also ends at hospital discharge. Kidney transplantation is identified with procedure or ICD-10 codes performed in the operation room. In-hospital death as recorded in Amsterdam UMC premises. While in the target trial, censoring occurs by design (e.g., administrative end of follow-up); in the emulation, censoring also happens due to hospital discharge, lack of monitorisation of serum creatinine, etc. In the target trial, follow-up is guaranteed until day 14, regardless of discharge status. In contrast, in the emulation, follow-up ends at discharge, which may shorten observation time and introduces the risk of informative censoring.
	Primary outcome	AKI within a 14-day risk window after exposure. Acute kidney injury was defined as per 2012 AKI KDIGO criteria using the serum creatinine criterion.	Same as target trial. Baseline serum creatinine is defined as the median serum creatinine between 7 to 365 days prior to hospital admission as suggested by Siew et al. or, in its default, first measurement during admission as suggested by Joyce et al.
	Causal contrasts	As treated or per-protocol (i.e., the effect of being assigned to a particular treatment strategy and complying with the protocolled treatment regimen); Intention-to-treat (i.e., the effect of being assigned to a particular treatment strategy, to estimate total effect)	Observational analogue of as treated or per-protocol effect (i.e., the effect of *initiating* a particular treatment strategy, and – once started – complying with the protocolled treatment regimen over the course of follow-up); Observational analogue of intention-to-treat effect (i.e., the effect of *initiating* a particular treatment strategy to estimate the total effect).
Identifying Assumptions		Intention-to-treat effect, total effect:	Conditional exchangeability holds given baseline covariates used in the IPTW models. Baseline confounders: demographic (i.e., sex, age),

Protocol Element	Target Trial Specification	Emulation with Observational Data
	Randomised treatment assignment: • Loss to follow-up: administrative censoring, hospital discharge, and 14 days after start of follow-up. • Intercurrent/competing events: receipt of kidney replacement therapy, out-of-class treatment strategy switch/add-on, documentation of pregnancy, prohibited medications, prolonged use of rescue medications, treatment discontinuation, in-hospital death, and drug-drug interactions that cause AKI.	healthcare utilisation (i.e., major surgery, bariatric surgery), comorbidities (i.e., prior AKI, cardiovascular disease, hypoalbuminaemia, sepsis, acute infections, chronic infections, Coronavirus disease 2019 [COVID-19], alcohol-related disorders, chronic kidney disease, liver disease, obesity, anaemia, chronic lung diseases, active cancer, diabetes mellitus, microangiopathics, vasculitides, renovascular disease, malignant hypertension, scleroderma, thrombosis of large arteries), medications (i.e., selected drugs with AKI potential, drug-drug interactions that cause AKI), mechanical (i.e., mechanical ventilation, benign prostatic hyperplasia, neurogenic bladder, intra/extra ureteric obstruction, retroperitoneal fibrosis, burns, trauma), genetic (i.e., glucose-6-phosphate dehydrogenase deficiency), vascular (i.e., hypovolaemia), ASA score, baseline serum creatinine, and rhabdomyolisis. The different detailed clinical models (DCMs) used for identifying confounders are further specified in the *Supplement*. Stable Unit Treatment Value Assignment (SUTVA) and Consistency: patient treatment does not affect the risk of another patient's AKI risk. No risk of multiple versions. Consistency (well-defined interventions), as drugs are identified with ATC codes, only systemic administration routes are considered, same index date for start of follow-up, and no multimodal analgesia. Positivity: non-zero probability of each treatment across covariate strata realised (i.e., covariate strata used in the weighting models) in the data. Independent censoring (IPCW): holds conditional on the observed covariate history and intercurrent events up to time *t*. Censoring is independent of potential outcomes given covariate history and intercurrent events up to time *t*. Correct model specification for the weighting/estimation

Protocol Element	Target Trial Specification	Emulation with Observational Data
		steps.
	Per protocol effect; controlled direct effect: Assumption of conditional exchangeability: same baseline confounders as per intention-to-treat effect.	Same as per intention-to-treat effect. Conditional exchangeability holds given baseline and time-varying covariates used in the IPTW models. Independent censoring (IPCW): holds conditional on the observed covariate history up to time *t*. Censoring is independent of potential outcomes given covariate history up to time *t*.
Estimator	Adjusted cumulative incidence curves were computed to visualise differences between groups, after applying inverse probability of censoring weighting (IPCW) to account for selection bias due to censoring (and intercurrent events), and inverse probability weighting (IPTW) to adjust for confounding, comparing metamizole with opioids and with NSAIDs. Propensity scores for IPCW were estimated using pooled logistic regression, and generalised propensity scores for IPTW were estimated using logistic regression. For intention-to-treat, total effect analysis, intercurrent events are allowed to occur, and are not censored on the basis of their occurrence. Stabilised IPCW are used to adjust for informative censoring due to loss to follow-up, where weights are constructed conditional on covariate history up to time *t* (including intercurrent event history). For the per protocol, controlled direct effect analysis, IPCW is applied both to censoring and intercurrent events to align with the per-protocol policy. Weights are constructed conditional on covariate history. Notably, a single-arm unstabilised weight was applied for the arm-specific potential drug-drug interactions (DDIs) to effectively eliminate the event, reflecting its near-zero probability in the remaining arms. For other censoring/intercurrent events, stabilised weights are used to reduce variance and to reflect the overall (marginal) distribution of covariates (i.e., how common each	As per target trial. Handling of missing data: single imputation with median for continuous variables and mode for categorical variables. For longitudinal variables, single imputation using forward filling. CKD patients were identified as per Fernández-Llaneza et al. *Clin Kidney J.* 2025; 18 (4).

Protocol Element	Target Trial Specification	Emulation with Observational Data
	covariate pattern is in the population). This helps ensure these events occur at random with respect to confounders and improves finite-sample performance. Subgroup analysis was carried out for CKD patients. Sensitivity analysis to assess dose-response relationship and inclusion of urine output criterion for AKI definition were performed.	
	Intention-to-treat estimands for total effect: risk difference, risk ratio. Utilisation of weighted Aalen-Johansen estimator.	
	Per-protocol estimands for controlled direct effect: risk difference, risk ratio. Utilisation of weighted Kaplan-Meier estimator.	

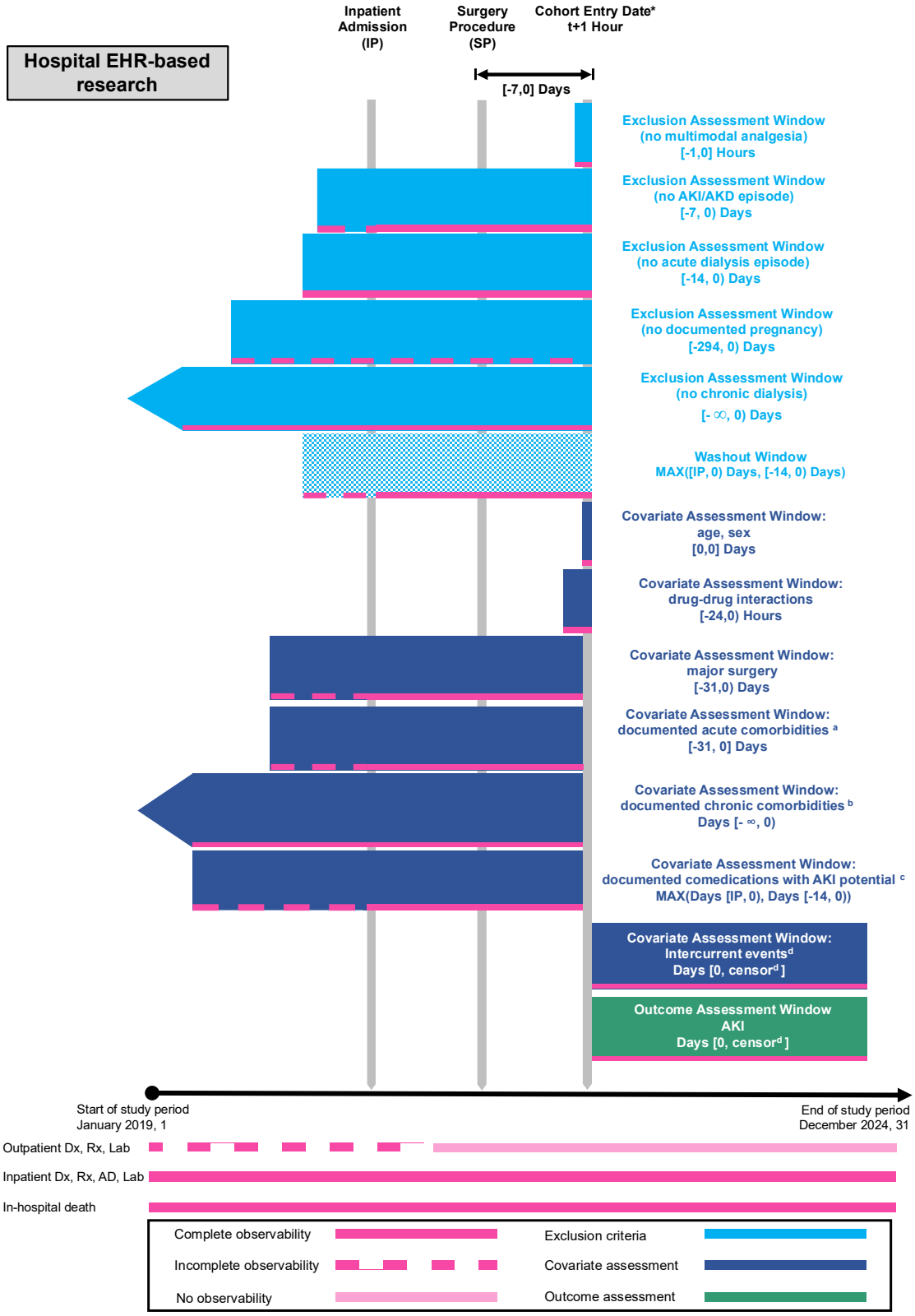
7. Research methods

7.1. Study design

Research design (e.g. cohort, case-control, etc.): New user active comparator cohort study

Rationale for study design choice: clear temporal relationship between exposure (initiation of metamizole) and outcome (i.e., incident AKI), incident exposure, and accurate risk estimation. This study design reduces risk of confounding by indication, prevalent user bias, and fits neatly into the target trial emulation framework.

7.2. Study design diagram



* t indicates treatment initiation

[a] Acute comorbidities: CVD (i.e., myocardial infarction), hypoalbuminaemia, sepsis, microangiopathics (HUS, TTP), vasculitides, acute infections, COVID-19, burns, trauma, hypovolaemia, and mechanical ventilation. Identified via ICD-10 or CBV codes listed in the *Supplement*. ASA score, and baseline serum creatinine identified from laboratory results.

[b] Chronic comorbidities/genetic disorders/procedures: CVD, chronic lung diseases, cancer, liver disease, renovascular disease, malignant hypertension, scleroderma, glucose-6-phosphate dehydrogenase deficiency, anaemia, BPH, neurogenic bladder, intra/extra-ureteric obstruction, retroperitoneal fibrosis, alcohol-related disorders, obesity, thrombosis of large arteries, and rhabdomyolisis identified via ICD-10 codes listed in the *Supplement*.

[c] Comedications with AKI potential with a frequency ≥1% in the SmPC: losartan, gentamicin, piperacillin and beta-lactamase inhibitor, vancomycin, amphotericin B, mitomycin, carboplatin, cisplatin, sirolimus, lithium, abacavir, cidofovir, foscarnet, ganciclovir, valganciclovir, zoledronic acid, hydrochlorothiazide, tacrolimus, ciclosporin). Drugs are identified via ATC codes.

[d] Intercurrent events: in-hospital death, receipt of kidney replacement therapy, identification of pregnancy, treatment arm switch/add-on, potential drug-drug interactions that cause AKI, prohibited medications (i.e., drugs with AKI potential), switch or add-on, administrative censoring or 14 days from treatment initiation.

Design based on framework from Wang SV, Schneeweiss S. A Framework for Visualizing Study Designs and Data Observability in Electronic Health Record Data. Clin Epidemiol. 2022 Apr 29;14:601-608. doi: 10.2147/CLEP.S358583

Figure 1. Pharmacoepidemiological design diagram

7.3. Setting

7.3.1 Context and rationale for definition of time 0 (and other primary time anchors) for entry to the study population

Time 0 is defined as one hour after pain relief inpatient medication administration (i.e., metamizole, opioids or NSAIDs) in a postoperative setting, this to allow a grace period in which intended multimodal therapy could be initiated.

Table 3 Operational Definition of Time 0 (index date) and other primary time anchors

Patients entered the base cohort based on initiation of pain relief medications (i.e., analgesics) like metamizole or a comparator drug (i.e., opioids or NSAIDs) after surgery. This allowed identification of patients in a postoperative pain management setting.

Study population name(s)	Time Anchor Description (e.g. time 0)	Number of entries	Type of entry	Washout window	Care Setting [1]	Code Type [2]	Diagnosis position	Incident with respect to…	Measurement characteristics / validation	Source of algorithm
Inpatients in postoperative pain management setting	Date of incident administration of pain relief medication (i.e., metamizole, NSAIDs, opioids)	Multiple	Incident	[-14,0] days	IP	n/a	n/a	Pain relief medication (systemic formulations)	No validation study	Investigator review of ATC and internal codes

[1] IP = inpatient admission, n/a = not applicable

[2]See appendix for listing of clinical codes for each study parameter

7.3.2 Context and rationale for study inclusion criteria:

All adult patients (≥18 years old) admitted to Amsterdam UMC for more than 24 hours were included as inpatients. Outpatient episodes in the same setting were considered for baseline covariates such as comorbidities and medication use

Table 4. Operational Definitions of Inclusion Criteria

Criterion	Details	Order of application	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
Inpatient admission	Trajectory ≥24 hours	Before selection of index date	[-∞,0]	IP	n/a	n/a	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	n/a	n/a
Adult	Age ≥18 years	Before selection of index date	[0,0]	IP	n/a	n/a	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	n/a	n/a
Postoperative pain management setting	All surgeries performed within Amsterdam UMC	Before selection of index date	[-7,0]	IP	n/a	n/a	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	n/a	n/a

[1] IP = inpatient admission, OP = outpatient episode, n/a = not applicable
[2] See appendix for listing of clinical codes for each study parameter
[3] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter)

7.3.3 Context and rationale for study exclusion criteria

We excluded the following patients: kidney transplant recipients, those on chronic dialysis, those who received acute dialysis, patients with recent AKI or AKD, individuals with prior use of the treatment medication, women documented as pregnant, and those on multimodal analgesia prior to baseline. Kidney transplant recipients were excluded due to potentially unstable kidney function. Patients on chronic dialysis and acute dialysis episodes fourteen days prior to time zero were excluded either because they may lack sufficient residual kidney function to sustain an injury or may still be recovering from unresolved kidney injury. Those with recent AKI or AKD were excluded to avoid including patients with prevalent elevations in serum creatinine. Pregnant women were excluded, as AKI during pregnancy is more likely to result from causes unrelated to the study drug. Multimodal analgesia patterns identified within the first hour of exposure administration are excluded, in order to assess patients exposed to exclusively one analgesic agent. It is considered that clinicians planned to give the medications simultaneously and the existing lag between medication administration and registrations is attributed to delays in the clinical setting. A 14-day washout period was applied to ensure new-user status.

Table 5. Operational Definitions of Exclusion Criteria

Criterion	Details	Order of application	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
Kidney transplantation		After selection of index date	[-365,0] days	IP/OP	ICD-10, CBV (local Dutch codes)	Any	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	No validation study	n/a
Chronic dialysis	Two dialysis measurements separated by 90 days with 2 dialysis measurements per week (or the equivalent of 25 encounters within that time window) or the registration of a relevant diagnosis code.	After selection of index date	[-∞,0]	IP/OP	ICD-10, CBV (local Dutch codes)	Any	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	No validation study	Fernández-Llaneza et al. *Clin Kidney J.* 2025; 18 (4)
Acute dialysis		After selection of index date	[-14,0] days	IP/OP	ICD-10, CBV (local Dutch codes)	Any	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	No validation study	Chawla et al. *Nat Rev Nephrol.* 2017; 13 (4):241-267
AKI or AKD episode	Derived from serum creatinine measurements. AKI: as per KDIGO guidelines, if serum	After selection of index date	[-7,0] days	IP/OP	n/a	n/a	Metamizole-exposed patients; Opioids-exposed	No validation study	KDIGO guidelines for AKI (2012); Levey. *Nephron*. 2022, 146 (3): 302-305;

Criterion	Details	Order of application	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
	creatinine increases by ≥26.5 µmol/L over 2 days or increased by 1.5-fold from baseline over 7 days. Recovery defined AKD: if the most recent eGFR to hospitalisation is less than 60 mL/min/1.73 m^2, and the preceding measure at least 3 months before is greater than or equal to 60 mL/min/1.73 m^2, or there was no preceding serum creatinine or eGFR measurement and albuminuria was absent or not measured prior to index date.						patients; NSAIDs-exposed patients.		James MT, et al. *JAMA Netw Open*. 2019, 2 (4):e191795
Documentation of Pregnancy	Pregnancy status	After selection of index date	[-294,0] days	IP/OP	n/a	n/a	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	No validation study	n/a
Multimodal analgesia (pain relief treatment switch/add-on)	Combination therapy which differs from any of the arms of the hypothetical trial.	After selection of index date	[-1,0] hours	IP	ATC, internal codes	n/a	Metamizole-exposed patients; Opioids-exposed patients; NSAIDs-exposed patients.	No validation study	Investigator review of medication names

[1] IP = inpatient, OP = outpatient, n/a = not applicable
[2] See appendix for listing of clinical codes for each study parameter
[3] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter)

7.4. Variables

7.4.1 Context and rationale for exposure(s) of interest

The exposure of interest is metamizole use, defined as first-ever inpatient administration occurring between surgery and up to seven days after surgery. A 14-day follow-up after initiation is used to allow sufficient time for the detection of AKI, ensure that all patients have a uniform observation period and reduce risk of immortal time bias.

Two active comparators are selected:

- Opioids, which are expected to have minimal AKI potential;
- Non-steroidal anti-inflammatory drugs (NSAIDs), which are known to increase AKI risk and serve as a relevant clinical comparator.

This exposure definition supports new-user, active-comparator design, minimising confounding by indication and aligning with real-world clinical practice.

Algorithm to define duration of exposure effect:

Exposure will be classified as first-ever use of metamizole, opioids, or NSAIDs in the postoperative period, as defined above. No stockpiling or refill-based assumptions are required, as this study focusses on inpatient medication records with medication administration timestamps.

Table 6. Operational Definition of Exposure

Exposure group name(s)	Details	Washout window	Assessment Window	Care Setting[1]	Code Type[2]	Diagnosis position[3]	Applied to study populations:	Incident with respect to...	Measurement characteristics/ validation	Source of algorithm
Exposure: metamizole	systemic administration	[-14,0)	[IP, SP]	IP	ATC, internal codes	n/a	Inpatients in postoperative pain management setting	Metamizole	No validation study	Investigator review of medication names
Comparator 1: opioids	systemic administration of buprenorphine, morphine, oxycodone, or piritramide	[-14,0)	[IP, SP]	IP	ATC, internal codes	n/a	Inpatients in postoperative pain management setting	Any medication in comparator 1	No validation study	Investigator review of medication names
Comparator 2: NSAIDs	systemic administration of aceclofenac, dexketoprofen, diclofenac, phenylbutazone, flurbiprofen, ibuprofen, indomethacin, ketoprofen, meloxicam, nabumetone,	[-14,0)	[IP, SP]	IP	ATC, internal codes	n/a	Inpatients in postoperative pain management setting	Any medication in comparator 2	No validation study	Investigator review of medication names

Exposure group name(s)	Details	Washout window	Assessment Window	Care Setting[1]	Code Type[2]	Diagnosis position[3]	Applied to study populations:	Incident with respect to...	Measurement characteristics/ validation	Source of algorithm
	naproxen, etoricoxib, parecoxib, celecoxib, piroxicam, propyphenazone, tiaprofenic acid, high-dose acetylsalycilic acid (≥100 mg/day) or diflusinal									

[1] IP = inpatient admission, SP = surgery procedure, n/a = not applicable
[2] See appendix for listing of clinical codes for each study parameter
[3] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter)

7.4.2 Context and rationale for outcome(s) of interest

The primary outcome of interest is AKI occurring during hospitalisation, defined according to the Kidney Disease: Improving Global Outcomes (KDIGO) criteria from 2012, based on changes in serum creatinine and/or urine output. This outcome is clinically meaningful, well-standardised, and captured in hospital electronic health records from Amsterdam UMC, where outpatient and inpatient data is routinely collected.

Acute kidney injury is a frequent and serious complication in postoperative patients, which is associated with increased morbidity, long-term damage of kidney function, and prolonged hospitalisations. Non-steroidal anti-inflammatory agents are known to contribute to AKI through kidney damage and dysfunction. On the other hand, opioids are not known to cause AKI, but they have been associated with respiratory depression and addiction as adverse drug events. Metamizole (dipyrone) is an alternative non-opioid analgesic often thought to have a different kidney safety profile.

Thus, outcome was selected to address a clinically and pharmacologically relevant safety question. This is important given the debate regarding the balance between analgesic efficacy and kidney safety in the postoperative setting.

Table 7. Operational Definitions of Outcome

Outcome name	Details	Primary outcome?	Type of outcome	Washout window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source of algorithm
Acute kidney injury	Using serum creatinine measurements only as recorded in laboratory	Yes	binary	[-7,0) days	IP	n/a	n/a	Inpatients in postoperative pain management setting	Variable depending on setting. Most studies conducted in emergency	KDIGO (2012) guideline for AKI; Siew et al. 2012. 7 (5): 712-9 (baseline serum creatinine)

Outcome name	Details	Primary outcome?	Type of outcome	Washout window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source of algorithm
	measurements. Baseline serum creatinine will be defined as 7 to 365 days prior to admission.								settings or ICU. 68% PPV, 79% sensitivity, 94% specificity 94% (Jonsson et al. Eur J Intern Med. 2019 60:78-82). Other examples in Hall PS et al. The future for diagnostic tests of acute kidney injury in critical care: evidence synthesis, care pathway analysis and research prioritisation. Health Technol Assess. 2018 22 (32):1-274	

[1] IP = inpatient, OP = outpatient, ED = emergency department, OT = other, n/a = not applicable
[2] See appendix for listing of clinical codes for each study parameter
[3] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter)

7.4.3 Context and rationale for follow up

The assessment window for the outcome of AKI is 14 days following first-ever inpatient administration of a pain relief medication (i.e., metamizole, opioids or NSAIDs). Acute kidney injury tends to develop in 7 days or less and there is a time-lag between kidney function and serum creatinine of around 2 days. Fourteen days are allowed to detect the outcome, capturing both acute onset and delayed presentations (e.g., cumulative effects).

The lookback window is of 1 year prior to treatment initiation, to capture healthcare utilisation potential confounders.

Table 8. Operational Definitions of Follow Up

Controlled Direct Effect

Follow up start	Day 1	
Follow up end[1]	**Select all that apply**	**Specify**
Date of outcome	Yes	Acute Kidney Injury
Date of death	Yes	In-hospital death
Other intercurrent events	Yes	start of kidney replacement therapy, drug-drug interactions
End of observation in data	Yes	End of study period.
Day X following index date *(specify day)*	Yes	Day 14
End of study period (specify date)	Yes	December 31, 2024
End of exposure *(specify operational details, e.g. stockpiling algorithm, grace period)*	No	In-class treatment discontinuation is disregarded; follow-up continues.
Date of add to/switch from exposure *(specify algorithm)*	Yes	switch/add-on to alternative pain relief medication
Other date *(specify)*	n/a	

[1] Follow up ends at the first occurrence of any of the selected criteria that end follow up.

Total Effect

Follow up start	Day 1	
Follow up end[1]	**Select all that apply**	**Specify**
Date of outcome	Yes	Acute Kidney Injury
Date of death	Yes	In-hospital death
Other intercurrent events	Yes	start of kidney replacement therapy, drug-drug interactions
End of observation in data	Yes	end of study period
Day X following index date *(specify day)*	Yes	Day 14

End of study period (specify date)	Yes	December 31, 2024
End of exposure *(specify operational details, e.g. stockpiling algorithm, grace period)*	No	In-class treatment discontinuation is disregarded; follow-up continues, because it preserves the total effect of the initial treatment strategy.
Date of add to/switch from exposure *(specify algorithm)*	Yes	Switch to alternative pain relief medication
Other date *(specify)*	n/a	

7.4.4 Context and rationale for covariates (confounding variables and effect modifiers, e.g. risk factors, comorbidities, comedications)

We included indications and contraindications for metamizole use which are known to cause AKI based on the Informatorium Medicamentorum curated and maintained by the Royal Dutch Pharmacists Association (Koninklijke Nederlandse Maatschappij ter bevordering der Pharmacie), further referred as IM-KNMP, a clinical nephrology textbook, Het Acute Boekje, and information from Yasrebi-de Kom et al. In addition, direct risk factors for AKI (i.e., causes of the outcome) were included, as those may also influence the treatment decision when AKI is suspected. For this, we consulted and supplemented were . Thus, comorbidities, comedications, healthcare utilisation, and social demographic covariates were considered. Variable selection We applied a prevalence-based variable selection technique. Selection of variables with a very low prevalence is acceptable as the risk of residual confounding is relatively small in such cases as per Patrick et al. Variables with binary values and associated prevalences below 1% were dropped.

Table 9. Operational Definitions of Covariates

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
age	at cohort entry	continuous	[0,0] days	n/a	n/a	n/a	Inpatients in postoperative pain management setting	n/a	n/a
sex	male, female	binary	[0,0] days	n/a	n/a	n/a	Inpatients in postoperative pain management setting	n/a	n/a
major surgery	CCI>3 AND (post-surgery ICU up to 12 hours after surgery OR operative time>4 h)	binary	[-31,0) days	ED, IP	n/a	n/a	Inpatients in postoperative pain management setting	No validation study	Investigator defined

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
bariatric surgery		binary	[SP,SP] days	IP	CBV	n/a	Inpatients in postoperative pain management setting	No validation study	Investigator defined
prior AKI	Defined as per 2012 KDIGO AKI guidelines using serum creatinine measurements	binary	[-365.25,-7] days	IP	internal codes	n/a	Inpatients in postoperative pain management setting	n/a	2012 KDIGO AKI
cardiovascular disease (chronic)	peripheral vascular disease, heart failure, arrhythmias	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Cardiac arrhythmias. Sensitivity: 39.0%, PPV: 93.4%, specificity: 99.2%, NPV: 85.3% Peripheral vascular disease. Sensitivity: 43.3%, PPV: 65.5%, specificity: 99.0%, NPV: 99.0%. Heart Failure. Sensitivity: 68.6%, PPV: 90.2%, specificity: 99.3%, NPV: 97.2% (Quan et al. 2008. Health Serv Res. 43 (4))	Quan et al. 2005. Med Care. 43 (11): 1130-9
cardiovascular disease (acute)	myocardial infarction	binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	sensitivity: 61.5%, PPV: 93.5%, specificity: 99.4%, NPV:	Quan et al. 2005. Med Care. 43 (11): 1130-9

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
								94.6% (Quan et al. 2008. Health Serv Res. 43 (4))	
chronic lung diseases	asthma, chronic obstructive pulmonary disease	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	sensitivity: 52.8%, PPV: 90.8%, specificity: 99.1%, NPV: 92.2% (Quan et al. 2008. Health Serv Res. 43 (4))	Quan et al. 2005. Med Care. 43 (11): 1130-9
anaemia	Blood loss anaemia, deficiency anaemia	binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Blood loss anaemia. Sensitivity: 17.8%, PPV: 32.0%, specificity: 99.6%, NPV: 99.1% (Quan et al. 2008. Health Serv Res. 43 (4))	Quan et al. 2005. Med Care. 43 (11): 1130-9
active cancer		binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting		Quan et al. 2005. Med Care. 43 (11): 1130-9
hypoalbuminaemia	Defined as albumin blood levels <30 g/L	binary	[-31,0) days	IP, OP	n/a	n/a	Inpatients in postoperative pain management setting	Unknown	Laboratory results
sepsis		binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Sensitivity: 71.9%, Specificity: 85.4%, PPV: 88.2%, NPV: 66.6%	Jolley RJ et al. 2015. BMJ Open. 5(12)

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
obesity	Defined as body mass index>30 kg/m^2	binary		IP, OP	(laboratory measurements)	Any	Inpatients in postoperative pain management setting	Unknown	Laboratory results
liver disease		binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Sensitivity: 40.6%, PPV: 85.4%, specificity: 99.6%, NPV: 96.9%. (Quan et al. 2008. Health Serv Res. 43 (4))	Quan et al. 2005. Med Care. 43 (11): 1130-9
renovascular disease	renovascular hypertension	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
malignant hypertension	Uncomplicated and complicated hypertension	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Sensitivity: 68.3%, PPV: 93.1%, specificity: 97.8%, NPV: 87.7% (Quan et al. 2008. Health Serv Res. 43 (4))	Quan et al. 2005. Med Care. 43 (11): 1130-9
scleroderma	systemic sclerosis	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
thrombosis of large arteries		binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
microangiopathics	Haemolytic uremic syndrome, thrombotic thrombocytopaenic purpura	binary	[-31, 0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management	No validation study	Investigator review of ICD-10 codes

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
							setting		
vasculitides	vasculitis	binary	[-31, 0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
COVID-19	Coronavirus disease 2019	binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
diabetes mellitus	With and without chronic complication	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Diabetes with chronic complication. Sensitivity: 59.1%, PPV: 63.1%, specificity: 99.0%, NPV: 98.9%. Diabetes without chronic complication. Sensitivity: 75.8%, PPV: 88.5%, specificity: 98.7%, NPV: 96.8%. (Quan et al. 2008. Health Serv Res. 43 (4))	Quan et al. 2005. Med Care. 43 (11): 1130-9
acute infections	Inflammatory diseases of the central nervous system, intestinal infectious diseases, zoonotic bacterial diseases, sexual infections,	binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
	spirochaetal diseases, rickettsioses, viral infections, mycoses, protozoal diseases, helminthiases, pediculosis, acariasis, sequelae of infectious and parasitic diseases, acute upper respiratory tract infections, influenza and pneumonia, infections of skin and subcutaneous tissue								
chronic infections	HIV/AIDS, herpes, Epstein-Barr virus, chronic hepatitis B and hepatitis C, toxoplasmosis, cytomegaly, Lyme disease, Q fever, endocarditis, schistosomiasis, tuberculosis	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	HIV/AIDS. Sensitivity: 41.7%, PPV: 100%, Specificity: 100%, specificity: 99.7%. (Quan et al. 2008. Health Serv Res. 43 (4)) Chronic hepatitis B. Sensitivity: 74.3%, PPV: 86.6%. Chronic hepatitis C. Sensitivity: 77.6%, PPV: 93.8%. (Kuang A et al. 2024. 72 (4): 202744)	Quan et al. 2005. Med Care. 43 (11): 1130-9; Kuang A et al. 2024. 72 (4): 202744; investigator review of ICD-10 codes
glucose-6-phospate dehydrogenase deficiency	Genetic disorder	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management	No validation study	Investigator review of ICD-10 codes

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
							setting		
benign prostatic hyperplasia		binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
neurogenic bladder		binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
intra/extra ureteric obstruction	hydronephrosis with ureteropelvic junction obstruction, hydronephrosis with ureteral stricture, hydronephrosis with renal and ureteral calculous obstruction	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
retroperitoneal fibrosis		binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	No validation study	Investigator review of ICD-10 codes
alcohol-related disorders	alcohol abuse	binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Sensitivity: 52.2%, PPV: 83.7%, specificity: 99.2%, NPV: 96.3% (Quan et al. 2008. Health Serv Res. 43 (4))	Quan et al. 2005. Med Care. 43 (11): 1130-9
chronic kidney disease		binary	[-∞,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Sensitivity: 78.8%, PPV: 64.3%, specificity: 98.2%, NPV: 99.1%. (Quan et al.	Quan et al. 2005. Med Care. 43 (11): 1130-9

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
								2008. Health Serv Res. 43 (4))	
trauma	Traumatic injury	binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Sensitivity: 69.8% PPV: 84.2% (Kuang et al. 2024. J Epidemiol Popul Health. 72(4):202744)	Kuang et al. 2024. J Epidemiol Popul Health. 72(4):202744
hypotension		binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting		Investigator review of ICD-10 codes
hypovolaemia		binary	[-31,0) days	IP, OP	ICD-10	n/a	Inpatients in postoperative pain management setting		Investigator review of ICD-10 codes
burns	burns in respiratory tract, exposure to smoke, fire, and flames, contact with heat and hot substances, exposure to electrical current and radiation	binary	[-31,0) days	IP, OP	ICD-10	Any	Inpatients in postoperative pain management setting	Sensitivity: 43%-93% Specificity: 86%-99% PPV: 49%-94% NPV: 85%-98% (Mason et al. 2017. Burns. 43(2): 258-264)	Mason et al. 2017. Burns. 43(2): 258-264.
mechanical ventilation		binary	[-31,0) days	IP	CBV (local Dutch codes)	n/a	Inpatients in postoperative pain management setting	No validation	Investigator review of procedure codes
baseline serum creatinine	Defined as the averaged outpatient serum creatinine between 7 to 365 days before admission	continuous	[IP-365,IP-7] days	IP, OP	(laboratory measurements)	n/a	Inpatients in postoperative pain management setting	Intraclass correlation coefficient with respesct to reference standard is 0.91 (95% confidence	Siew et al. 2012. 7 (5): 712-9

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
								interval, 0.88-0.92)	
acute kidney injury		binary	[-365,-7] days	IP, OP	(laboratory measurements)	n/a	Inpatients in postoperative pain management setting		KDIGO (2012) guideline for AKI
American Society of Anaesthesiology (ASA) score	Assesses the physical health status before surgery and anaesthesia with a scale from 1 (healthy) to 6 (brain-dead)	ordinal	[IP,0) days	IP	n/a	n/a	Inpatients in postoperative pain management setting	n/a	n/a
drugs with AKI potential – Agents acting on the renin-angiotensin-aldosterone system	losartan (frequent, 1-10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential – Antibacterials for systemic use	gentamicin (very frequent, >10%), piperacillin and beta-lactamase inhibitor (frequent 1-10%) vancomycin (frequent 1-10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential – antimycotics for systemic use	Amphotericin B (frequent 1-10%)	Binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential – Antineoplastic agents	mitomycin (frequent, 1-10%), carboplatin (very frequent, >10%), cisplatin (very frequent, >10%), sirolimus (frequent, 1-10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
drugs with AKI potential - Antipsychotics	lithium (very frequent, >10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential - Antivirals for systemic use	abacavir (very frequent, >10%), cidofovir (frequent, 1-10%), foscarnet (frequent, 1-10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential - Bisphosphonates	zoledronic acid (frequent, 1-10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential - Diuretics	hydrochlorothiazide (frequent, 1-10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential - Immunosuppressants	tacrolimus (frequent, 1-10%), ciclosporin (frequent, 1-10%)	binary	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with AKI potential	Directly derived from the sum of binary indicators for drugs with AKI potential (i.e., selected agents acting on the renin-angiotensin-aldosterone system, antibacterials, antimycotics, antineoplastic, antipsychotics, antivirals, bisphosphonates, diuretics, immunosuppressants	numerical (discrete)	max([-14, 0) days, [IP,0) days)	IP, OP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	No validation study

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
	)								
drugs with (drug-drug) interactions with metamizole that cause AKI	Twenty four hours are added to the last administration of the drug to construct treatment episodes and assess overlap with metamizole.	binary	Confounder: [-24,0) hours Intercurrent event: [0,14] days	IP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	IM-KNMP Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drugs with (drug-drug) interactions with NSAIDs that cause AKI	Twenty four hours are added to the last administration of the drug to construct treatment episodes and assess overlap with NSAIDs. These include interactions of NSAIDs with beta-blocking agents, agents acting on the renin-angiotensin system, and diuretics. Note that all single drugs from these pharmacological classes are derived from the ones indicated in IM-KNMP	binary	Confounder: [-24 hours,0) Intercurrent event: [0,14] days	IP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	IM-KNMP Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58
drug-drug interactions that cause AKI	Agents acting on the renin-angiotensin system and diuretics, tacrolimus and drugs with AKI potential, ciclosporin and drugs with AKI potential. They do not include any drug from the arms of the hypothetical trial. Note that all single drugs from these	binary	max([-14, 0) days, [IP,0) days) Intercurrent event: [0,14] days	IP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	IM-KNMP Fernández-Llaneza et al. 2025, Drug Saf. 48(1): 43-58

Characteristic	Details	Type of variable	Assessment window	Care Settings[1]	Code Type[2]	Diagnosis Position[3]	Applied to study populations:	Measurement characteristics/ validation	Source for algorithm
	pharmacological classes are derived from the ones indicated in IM-KNMP								
death	In-hospital death	binary	Intercurrent event: [0,14] days	IP	n/a	n/a	Inpatients in postoperative pain management setting	No validation study	
pregnancy		binary	Intercurrent event: [0,14] days	IP	n/a	n/a	Inpatients in postoperative pain management setting	No validation study	
dialysis		binary	Intercurrent event: [0,14] days	IP/OP	ICD-10, CBV (local Dutch codes)	Any	Inpatients in postoperative pain management setting	No validation study	Fernández-Llaneza et al. *Clin Kidney J.* 2025; 18 (4)
treatment switch/add-on	Change from one treatment arm to another	binary	Intercurrent event: [0,14] days	IP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	
Prolonged analgesic rescue medication	Applicable to paracetamol (acetaminophen), and tramadol only when administered for more than 48 hours (continuously).	binary	Intercurrent event: [0,14] days	IP	ATC	n/a	Inpatients in postoperative pain management setting	No validation study	

[1] IP = inpatient admission, SP = surgery procedure, OP = outpatient episode, ED = emergency department, OT = other, n/a = not applicable
[2] See appendix for listing of clinical codes for each study parameter
[3] Specify whether a diagnosis code is required to be in the primary position (main reason for encounter)

7.5. Data analysis

7.5.1 Context and rationale for analysis plan

We use multinomial logistic regression to estimate propensity scores to apply inverse probability weighting (IPTW) to adjust for confounding. Pooled logistic regression was used to estimate propensity scores to apply inverse probability of censoring weighting (IPCW) to adjust for selection bias and estimate effects not mediated by intercurrent events, where relevant. We apply weighted cumulative incidence estimators to estimate the cumulative incidence function (CIF) and then calculate risk difference (and risk ratio). The primary estimand of interest is the controlled direct effect.

Table 10. Primary, secondary, and subgroup analysis specification

A. Primary analysis

B. Hypothesis:	Metamizole use increases the risk of acute kidney injury where receipt of kidney replacement therapy, documentation of pregnancy, administration of drugs with AKI potential (i.e., prohibited medications) or drug combinations that cause AKI, administration of rescue medications for more than 48 hours, out-of-class treatment switch/add-ons, in-hospital death, treatment discontinuation and loss to follow-up occur at random [total effect]
Exposure contrast:	First-ever administration of metamizole vs. first-ever administration of opioids (comparator 1) First-ever administration of metamizole vs. first-ever administration of NSAIDs (comparator 2)
Outcome:	Acute kidney injury (using serum creatinine criterion only)
Analytic software:	R v4.2.2; Python 3.13.5
Model(s): *(provide details or code)*	Propensity score model: multinomial logistic regression modelling treatment assignment (i.e., metamizole, opioids, NSAIDs) as a function of baseline confounders. Stabilised inverse probability of treatment weights will be calculated from this model. Censoring model: a pooled logistic regression model will estimate the probability of remaining uncensored over time, conditional on treatment, relevant time-varying covariates, and intercurrent events. Stabilised inverse probability weights for censoring (IPCW) will be derived. Weighting: inverse probability weights for treatment and censoring (i.e., loss to follow-up) events will be combined multiplicatively to create composite weights. Outcome model: cumulative incidence curves will be generated for each treatment group using inverse probability weighting to adjust for confounding, and censoring. The estimator corresponds to a weighted Aalen-Johansen estimator that accounts for intercurrent events. Risk differences and risk ratios comparing cumulative incidences at 14 days will be calculated.
Confounding adjustment method	***Name method and provide relevant details, e.g. bivariate, multivariable, propensity score matching (specify matching algorithm ratio and caliper), propensity score weighting (specify weight formula, trimming, truncation), propensity score stratification (specify strata definition), other.***
	To adjust for baseline confounding, we will estimate generalised propensity scores using a multinomial logistic regression

	model. The probability of receiving each treatment will be modelled conditional on a set of baseline covariates, including sociodemographic characteristics, comorbidities, healthcare utilisation, and concomitant medications. Individuals with extreme propensity scores, defined as scores below 1st percentile or above the 99th percentile, will be excluded to improve overlap and reduce the influence of outliers. The resulting stabilised inverse probability of treatment weights will be used to balance covariates across treatment groups and estimate marginal treatment effects.
Missing data methods	***Name method and provide relevant details, e.g. missing indicators, complete case, last value carried forward, multiple imputation (specify model/variables), other.***
	Missing values will be imputed using the Multivariable Imputation by Chained Equations (MICE), assuming missing at random missingness distribution. Variables with >30% missingness will be excluded. Categorical variables should have a minor category proportion of at least 5%.
Subgroup Analyses	***List all subgroups***
	If sample size is sufficient, chronic Kidney Disease patients in Amsterdam UMC, as defined by Fernández-Llaneza et al. 2025, CKJ, 18(4), will be studied. CKD stage at administration date will be introduced as a confounder.

C. Secondary analysis

Secondary analysis 1

Hypothesis:	Metamizole use increases the risk of acute kidney injury in the hypothetical setting where receipt of kidney replacement therapy, documentation of pregnancy, administration of drugs with AKI potential (i.e., prohibited medications) and drug-drug interactions that cause AKI, use of rescue medication for more than 48 hours, out-of-class treatment switches/add-ons, in-hospital death, treatment discontinuation occurs at random and treatment-specific potential drug-drug interactions do not occur. [controlled direct effect]
Exposure contrast:	First-ever administration of metamizole vs. first-ever administration of opioids (comparator 1) First-ever administration of metamizole vs. first-ever administration of NSAIDs (comparator 2)
Outcome:	Acute kidney injury
Analytic software:	R v4.2.2; Python 3.13.5
Model(s): ***(provide details or code)***	Propensity score model: multinomial logistic regression, with the treatment (i.e., metamizole, NSAIDs, opioids) as the dependent variable and the confounders as independent variables. The selected confounders are those which affect both the exposure and the outcome or ' mediator' (i.e., intercurrent event) and outcome relationship. Stabilised inverse probability of treatment weights will be calculated from this model. Censoring and intercurrent event models: pooled logistic regression models to estimate the probabilities of remaining uncensored and free of intercurrent events over time, conditional on treatment and relevant time-varying covariates. Stabilised inverse probability weights for censoring and intercurrent events will be derived, except for arm-specific drug-drug interactions, whereby unstabilised inverse probability weights for censoring will be derived to effectively eliminate censoring in the pseudo-population. Weighting: inverse probability weights for treatment and censoring will be combined multiplicatively to create composite weights.

	Outcome model: cumulative incidence curves will be generated for each treatment group using inverse probability weighting with the composite weights to adjust for confounding, censoring, and intercurrent events. For the controlled direct effect, the estimator corresponds to the complement of a weighted Kaplan-Meier estimator. The risk difference and risk ratio comparing cumulative incidences at 14 days will be calculated.
Confounding adjustment method	***Name method and provide relevant details, e.g. bivariate, multivariable, propensity score matching (specify matching algorithm ratio and caliper), propensity score weighting (specify weight formula, trimming, truncation), propensity score stratification (specify strata definition), other.***
	To adjust for baseline confounding, we will estimate generalised propensity scores using a multinomial logistic regression model. The probability of receiving each treatment will be modelled conditional on a set of baseline covariates, including sociodemographic characteristics, comorbidities, healthcare utilisation, and concomitant medications. Individuals with extreme propensity scores, defined as scores below 1st percentile or above the 99th percentile, will be excluded to improve overlap and reduce the influence of outliers. The resulting stabilised inverse probability of treatment weights will be used to balance covariates across treatment groups and estimate marginal treatment effects.
Missing data methods	***Name method and provide relevant details, e.g. missing indicators, complete case, last value carried forward, multiple imputation (specify model/variables), other.***
	Missing values will be imputed using the Multivariable Imputation by Chained Equations (MICE), assuming missing at random missingness distribution. Variables with >30% missingness will be excluded. Categorical variables should have a minor category proportion of at least 5%
Subgroup Analyses	***List all subgroups***
	If sample size is sufficient, chronic Kidney Disease patients in Amsterdam UMC, as defined by Fernández-Llaneza et al. 2025, CKJ, 18(4), will be studied. CKD stage at administration date will be introduced as a confounder.

Secondary analysis 2

Hypothesis:	Metamizole use increases the risk of acute kidney injury regardless of receipt of kidney replacement therapy, documentation of pregnancy, administration of drugs with AKI potential or drug combinations that cause AKI, administration of rescue medications for more than 48 hours, treatment switch/add-on, in-hospital death, or treatment discontinuation [total effect]
Exposure contrast:	First-ever administration of metamizole vs. first-ever administration of opioids (comparator 1) First-ever administration of metamizole vs. first-ever administration of NSAIDs (comparator 2)
Outcome:	Acute kidney injury (using serum creatinine criterion only)
Analytic software:	R v4.2.2; Python 3.13.5
Model(s): ***(provide details or code)***	Propensity score model: machine learning model (e.g., XGBoost, SuperLearner) for treatment assignment (i.e., metamizole, opioids, NSAIDs) as a function of baseline confounders. Stabilised inverse probability of treatment weights will be calculated from this model. McCaffrey et al. and Piracchio et al. provide a good starting foundation for carrying these analyses. Censoring model: a pooled logistic regression model will estimate the probability of remaining uncensored over time, conditional on treatment, relevant time-varying covariates, and intercurrent events. Stabilised inverse probability weights

	for censoring (IPCW) will be derived. Weighting: inverse probability weights for treatment and censoring will be combined multiplicatively to create composite stabilised weights. Outcome model: cumulative incidence curves will be generated for each treatment group using inverse probability weighting to adjust for confounding, and censoring. The estimator corresponds to a weighted Aalen-Johansen estimator that accounts for intercurrent events. Risk differences and risk ratios comparing cumulative incidences at 14 days will be calculated.
Confounding adjustment method	***Name method and provide relevant details, e.g. bivariate, multivariable, propensity score matching (specify matching algorithm ratio and caliper), propensity score weighting (specify weight formula, trimming, truncation), propensity score stratification (specify strata definition), other.***
	To adjust for baseline confounding, we will estimate generalised propensity scores using a machine learning (i.e., non-parametric) model. The probability of receiving each treatment will be modelled conditional on a set of baseline covariates, including sociodemographic characteristics, comorbidities, healthcare utilisation, and concomitant medications. Individuals with extreme propensity scores, defined as scores below 1st percentile or above the 99th percentile, will be excluded to improve overlap and reduce the influence of outliers. The resulting stabilised inverse probability of treatment weights will be used to balance covariates across treatment groups and estimate marginal treatment effects.
Missing data methods	***Name method and provide relevant details, e.g. missing indicators, complete case, last value carried forward, multiple imputation (specify model/variables), other.***
	Missing values will be imputed using the Multivariable Imputation by Chained Equations (MICE), assuming missing at random missingness distribution. Variables with >30% missingness will be excluded. Categorical variables should have a minor category proportion of at least 5%.
Subgroup Analyses	***List all subgroups***
	N/A

Secondary analysis 3

Hypothesis:	Metamizole use increases the risk of acute kidney injury regardless of receipt of kidney replacement therapy, documentation of pregnancy, administration of drugs with AKI potential or drug combinations that cause AKI, administration of rescue medications for more than 48 hours, treatment switch/add-on, in-hospital death, or treatment discontinuation [total effect]
Exposure contrast:	First-ever administration of metamizole vs. first-ever administration of opioids (comparator 1) First-ever administration of metamizole vs. first-ever administration of NSAIDs (comparator 2)
Outcome:	Acute kidney injury (using serum creatinine and urine output criteria from KDIGO guidelines on AKI)
Analytic software:	R v4.2.2; Python 3.13.5
Model(s): ***(provide details or code)***	Propensity score model: multinomial logistic regression modelling treatment assignment (i.e., metamizole, opioids, NSAIDs) as a function of baseline confounders. Stabilised inverse probability of treatment weights will be calculated from this model.

	Censoring model: a pooled logistic regression model will estimate the probability of remaining uncensored over time, conditional on treatment, relevant time-varying covariates, and intercurrent events. Stabilised inverse probability weights for censoring (IPCW) will be derived. Weighting: inverse probability weights for treatment and censoring will be combined multiplicatively to create composite stabilised weights. Outcome model: cumulative incidence curves will be generated for each treatment group using inverse probability weighting to adjust for confounding, and censoring. The estimator corresponds to a weighted Aalen-Johansen estimator that accounts for intercurrent events. Risk differences and risk ratios comparing cumulative incidences at 14 days will be calculated.
Confounding adjustment method	***Name method and provide relevant details, e.g. bivariate, multivariable, propensity score matching (specify matching algorithm ratio and caliper), propensity score weighting (specify weight formula, trimming, truncation), propensity score stratification (specify strata definition), other.***
	To adjust for baseline confounding, we will estimate generalised propensity scores using a multinomial logistic regression model. The probability of receiving each treatment will be modelled conditional on a set of baseline covariates, including sociodemographic characteristics, comorbidities, healthcare utilisation, and concomitant medications. Individuals with extreme propensity scores, defined as scores below 1st percentile or above the 99th percentile, will be excluded to improve overlap and reduce the influence of outliers. The resulting stabilised inverse probability of treatment weights will be used to balance covariates across treatment groups and estimate marginal treatment effects.
Missing data methods	***Name method and provide relevant details, e.g. missing indicators, complete case, last value carried forward, multiple imputation (specify model/variables), other.***
	Missing values will be imputed using the Multivariable Imputation by Chained Equations (MICE), assuming missing at random missingness distribution. Variables with >30% missingness will be excluded. Categorical variables should have a minor category proportion of at least 5%.
Subgroup Analyses	***List all subgroups***
	N/A

Secondary analysis 4

Hypothesis:	Metamizole use increases the risk of acute kidney injury regardless of receipt of kidney replacement therapy, documentation of pregnancy, administration of drugs with AKI potential or drug combinations that cause AKI, administration of rescue medications for more than 48 hours, treatment switch/add-on, in-hospital death, or treatment discontinuation [total effect]
Exposure contrast:	First-ever administration of metamizole vs first-ever administration of diclofenac
Outcome:	Acute kidney injury
Analytic software:	R v4.2.2; Python 3.13.5
Model(s): ***(provide details or code)***	Propensity score model: logistic regression modelling treatment assignment (metamizole vs. diclofenac) as a function of baseline confounders. Stabilised inverse probability of treatment weights will be calculated from this model.

	Censoring model: a pooled logistic regression model will estimate the probability of remaining uncensored over time, conditional on treatment, relevant time-varying covariates, and intercurrent events. Stabilised inverse probability weights for censoring (IPCW) will be derived. Weighting: inverse probability weights for treatment and censoring will be combined multiplicatively to create composite stabilised weights. Outcome model: cumulative incidence curves will be generated for each treatment group using inverse probability weighting to adjust for confounding, and censoring. The estimator corresponds to a weighted Aalen-Johansen estimator that accounts for intercurrent events. Risk differences and risk ratios comparing cumulative incidences at 14 days will be calculated.
Confounding adjustment method	***Name method and provide relevant details, e.g. bivariate, multivariable, propensity score matching (specify matching algorithm ratio and caliper), propensity score weighting (specify weight formula, trimming, truncation), propensity score stratification (specify strata definition), other.***
	To adjust for baseline confounding, we will estimate generalised propensity scores using a logistic regression model. The probability of receiving each treatment will be modelled conditional on a set of baseline covariates, including sociodemographic characteristics, comorbidities, healthcare utilisation, and concomitant medications. Individuals with extreme propensity scores, defined as scores below 1st percentile or above the 99th percentile, will be excluded to improve overlap and reduce the influence of outliers. The resulting stabilised inverse probability of treatment weights will be used to balance covariates across treatment groups and estimate marginal treatment effects.
Missing data methods	***Name method and provide relevant details, e.g. missing indicators, complete case, last value carried forward, multiple imputation (specify model/variables), other.***
	Missing values will be imputed using the Multivariable Imputation by Chained Equations (MICE), assuming missing at random missingness distribution. Variables with >30% missingness will be excluded. Categorical variables should have a minor category proportion of at least 5%.
Subgroup Analyses	***List all subgroups***
	N/A

Table 11. Sensitivity analyses – rationale, strengths and limitations

	What is being varied? How?	Why? (What do you expect to learn?)	Strengths of the sensitivity analysis compared to the primary	Limitations of the sensitivity analysis compared to the primary
Sensitivity Analysis 1 [total effect]	Metamizole exposure definition; comparing high dose (i.e., ≥3,000 mg/day) vs. low dose (i.e., <3,000 mg/day) as per Brinkman et al. 2025. Br J Clin Pharmacol; 1-8, instead of metamizole vs. opioids.	To test the consistency assumption and address whether a dose-response relationship exists between metamizole and AKI.	Increases granularity of the exposure definition; provides insight into a possible dose-response, supporting causal interpretation.	No universally accepted thresholds, increasing risk of exposure misclassification; smaller sample sizes within dose strata may reduce precision

7.6. Data sources

7.6.1 Context and rationale for data sources

Reason for selection: The Research Data Platform (RDP) contains data from electronic medical records (Epic) from Amsterdam UMC. Amsterdam UMC consists of two teaching hospitals in Amsterdam. The whole hospital trajectory for admitted patients is fully traceable and there is also availability of data related to outpatient specialist clinics.

Strengths of data source(s): The database contains longitudinal records from secondary care, where clinicians and nurses enter information using Epic's standardised electronic forms. All prescriptions and medication administrations from clinicians are recorded, medical diagnoses and procedures are captured via ICD-10 and CBV codes, laboratory measurements, medical history, lifestyle/clinical variables such as body mass index are also captured. After entry in the system, all diagnoses codes are curated by specialised medical coders.

Limitations of data source(s): The data source reflects inpatient and outpatient encounters captured in electronic health records in Amsterdam UMC and may miss elements of care received in primary care or other institutions.

Data source provenance/curation: Access to the data source is provided after authorisation by the Research Data Management board and is validated periodically by medical coders. The selected data source is used for research and documentation of data contents is provided.

Table 12. Metadata about data sources and software

	Data 1
Data Source(s):	Amsterdam UMC Electronic Health Records
Study Period:	2019 – 2024
Eligible Cohort Entry Period:	2019 – 2024
Data Version (or date of last update):	16 March 2024
Data sampling/extraction criteria:	via Research Data Platform
Type(s) of data:	Electronic health records
Data linkage:	No external data linkage
Conversion to CDM*:	n/a
Software for data management:	RStudio 1.4.1106 and SQL Server Management Studio v19.2.56.2

*CDM = Common Data Model

7.7. Data management

Pseudonymised routinely collected EHR data of patients admitted to Amsterdam UMC were re-used. Laboratory findings were extracted from two laboratory information systems: GLIMS and LabTrain. All data was extracted retrospectively on 31st December 2023 covering hospital admissions between January 2019, 1 and December 2023, 31. Amsterdam UMC uses EHR system of EPIC© Epic Electronic Health Record [Computer software]. , *Madison, WI: Epic; Madison, WI*. Data extractions from the EHR system are facilitated by Research Data Management team of Amsterdam UMC via Research Data Platform and according to a Standard Operating Procedure "*Reuse of care data for the purpose of research*". EHR data are organized in tables called Detailed Clinical Models (DCMs). Each DCM contains variables compiled from various EHR tables, for example a DCM Problem List contains all registered problems per patient and admission linked to ICD-10 or ICD-9, as well as among others the status of the problem, date of registration, the provider number, location in hospital.

7.8. Quality control

This protocol was drafted in collaboration with experts in pharmacology, causality, pharmacoepidemiology, and medical informatics. Weekly meetings are held to discuss important topics on this study and potential updates of the study design and methodological challenges. Quality control on programming was checked by structuring the code independently, running and explaining it to peers and colleagues within the team.

7.9. Study size and feasibility

A preliminary feasibility counts study (see **Figure 2**) revealed a sample size of approximately 37,500 patients administered metamizole, 14,500 patients administered NSAIDs, and 50,000 patients administered opioids in a postoperative setting in Amsterdam UMC. After implementing preliminary exclusion/inclusion criteria, the counts get down to approximately 12,000 patients exposed to metamizole, 580 patients exposed to NSAIDs, and 14,000 patients exposed to opioids. When using serum creatinine as a criterion for identifying AKI following the KDIGO (2012) guideline, there are 1,022 AKI cases within 14 days of opioid initiation, 238 AKI cases within 14 days of metamizole initiation, and 14 cases within 14 days of NSAIDs initiation, respectively. The distribution of first AKI across the principal stratum (i.e., only patients who suffered an AKI episode) seems comparable across arms (see **Figure 3**).

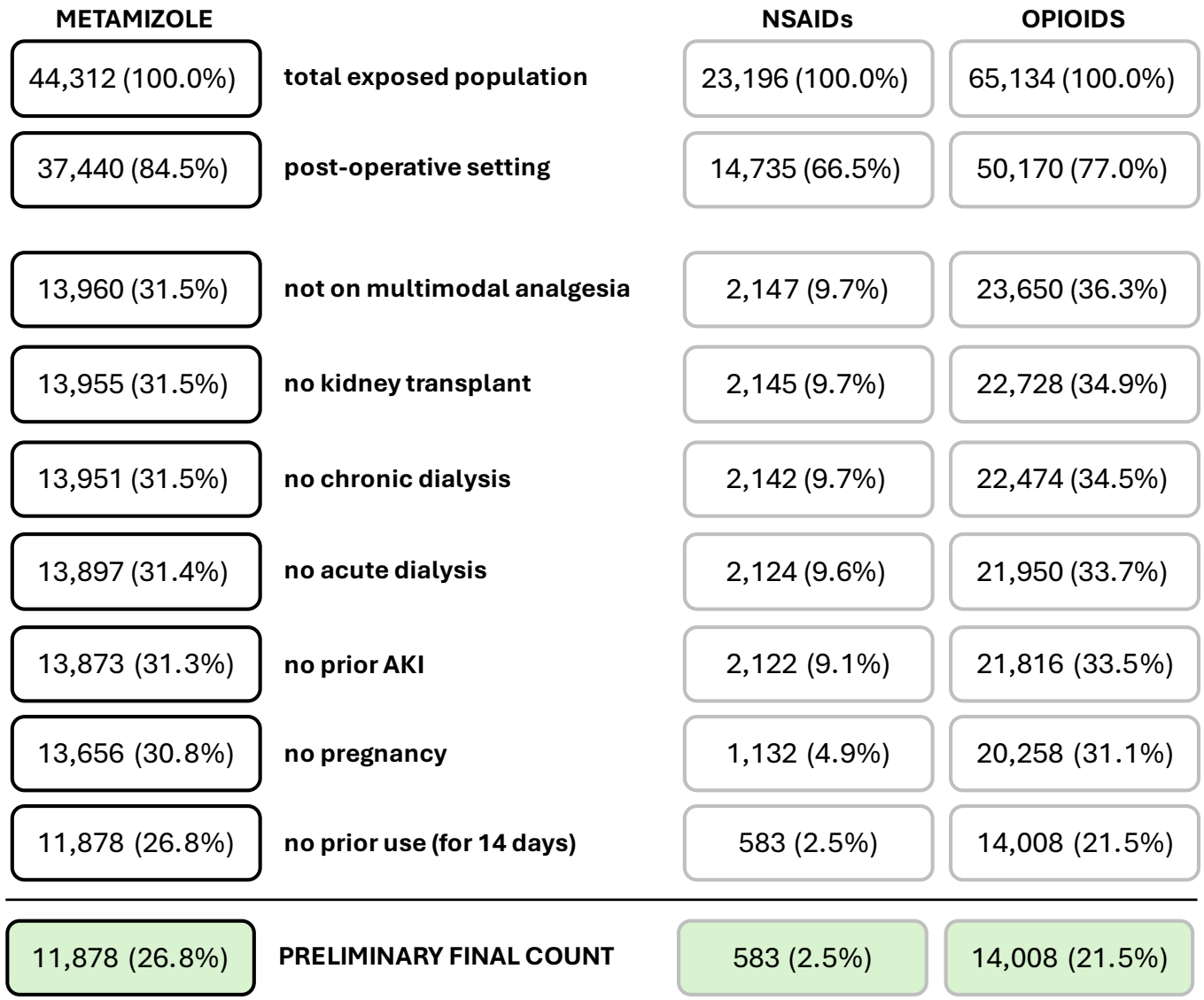


Figure 2. Preliminary attrition funnel for the three treatment arms in the hypothetical trial.

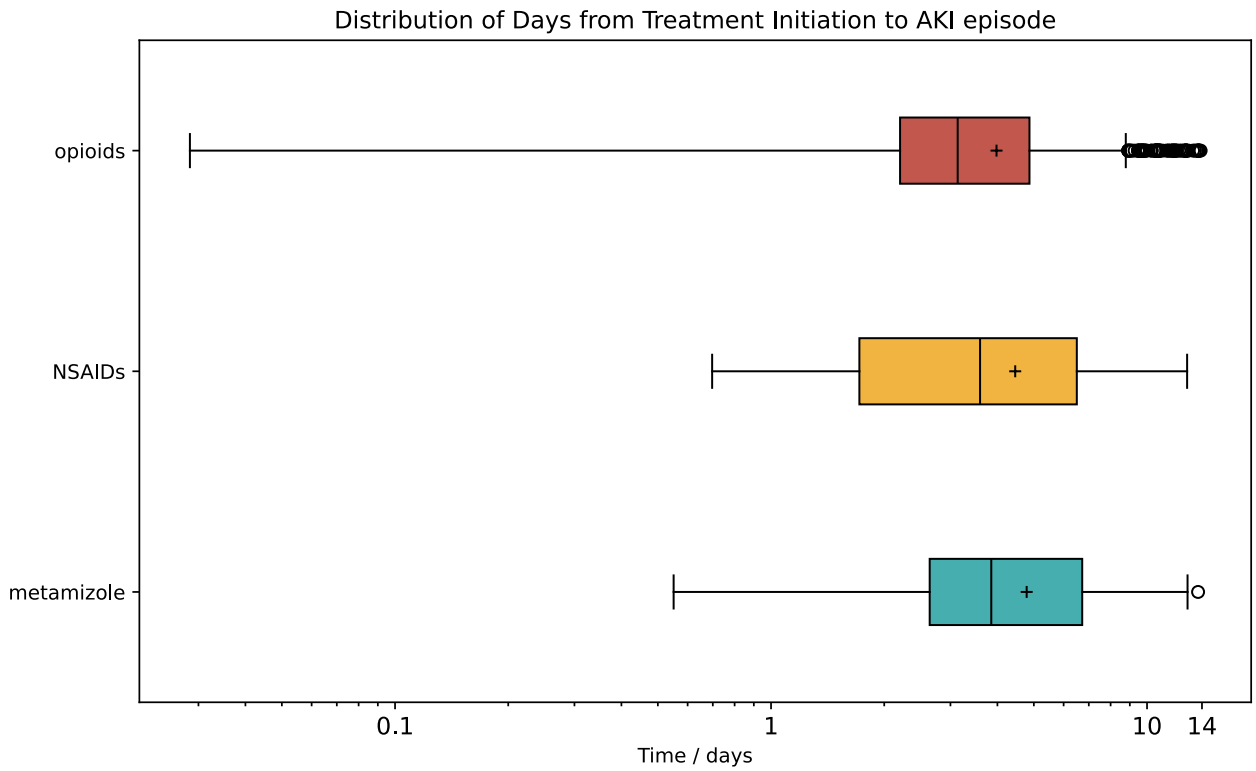


Figure 3. Boxplot displaying time distribution of first ever AKI event per treatment arm after time zero.

As previously mentioned, existing evidence on metamizole AKI potential is scarce and there are few clinical trials and observational studies clearly reporting on this. Stueber et al. report an odds ratio of 1.62 (95% CI: 1.12-2.34) for dose per day (i.e., a 1.6-fold increase in the incidence of AKI per each additional gram of intravenous metamizole administered per day). Baseline risk for AKI (i.e., patients not exposed to metamizole) is estimated to be around 20% as per Susantitaphong et al. Following the formula from Zhang et al., we can convert the odds ratio to a risk ratio and risk difference:

$$RR = \frac{OR}{(1 - p_0) + (p_0 \times OR)} = \frac{1.62}{(1 - 0.2) + (0.2 \times 1.62)} \approx 1.44\ (95\%\ CI: 1.09 - 1.85)$$

$$RD = 0.20 \times (1.44 - 1) = 0.124\ (12.4\%\ absolute\ risk\ increase)$$

To calculate the sample size, given baseline AKI risk of p_0=0.20, p_1=RR*p_0=1.44*0.20=0.29, power (i.e., the probability of correctly rejecting the null when the alternative is true) of 0.80 and significance level (α) of 0.05:

$z_{1-\alpha/2}$=2.241; $z_{1-\beta}$=0.84

compute pooled mean $\bar{p} = \frac{0.20+0.29}{2} = 0.245$

For a binary outcome within a fixed window (e.g., 14 days), assuming equal allocation, the standard two-sample approximate formula is:

$$n \approx \frac{\left(z_{1-\frac{\alpha}{2}}\sqrt{2\bar{p}(1-\bar{p})} + z_{1-\beta}\sqrt{p_1(1-p_1)+p_0(1-p_0)}\right)^2}{(p_1-p_0)^2} = \frac{\left(2.241 \times \sqrt{2 \times 0.245 \times 0.755} + 0.84 \times \sqrt{0.29 \times 0.71 + 0.20 \times 0.80}\right)^2}{(0.09)^2} \cong 432\ patients$$

About 432 patients would be needed per arm for each metamizole vs comparator comparison under α'=0.05 and 80% power. With three arms and equal allocation, a total of N = 432 x 3 = 1,269 patients would be required. This might be a conservative approach given that a Bonferroni correction has been applied. Under the stated assumptions, the study would be designed with approximately 1,296 participants (432 per arm). Thus, the study appears to be sufficiently powered, even under conservative scenarios.

Table 13. Distribution of single-drugs in treatment arms from hypothetical trial

Arm	Single drug	Unique patient counts after inclusion/exclusion criteria	Number of AKI cases* (based on SCr criterion only)	Crude incidence (%)
METAMIZOLE		**11,812**	**238**	**2.0**
NSAIDs	diclofenac	521	12	2.3
	ibuprofen	14	0	0
	naproxen	7	0	0
	celecoxib	6	0	0
	etoricoxib	5	1	20.0
	indometacin	1	0	0
	high-dose acetylsalicylic acid	12	4	33.3
	Grand total	**566**	**16**	**2.8**
OPIOIDS	morphine	8,372	781	9.3
	piritramide	5,041	217	4.3
	oxycodone	267	9	3.4
	buprenorphine	92	13	14.1
	Grand total	**13,772**	**1,020**	**7.4**

*without accounting for intercurrent events

8. Limitation of the methods

There are several potential limitations with the methods specified in this protocol and we outline some remediating strategies that we implemented.

1. The data was not collected for research and some important variables may not be collected or may be measured imperfectly.
    a. We have selected validated algorithms when possible
    b. We have considered outpatient data and medical history data for covariate assessment to reduce misclassification bias
    c. Serum creatinine measurements will be measured opportunistically rather than being systematically monitored
    d. We have designed a medication administration observation window of one hour to capture delay in registration of multimodal analgesia patterns (i.e., administration of more than one pain relief medication)
2. There will be no randomisation
    a. We have emulated the design of a target trial and present the estimands framework
    b. We have balanced compared groups with respect to confounding variables via IPTW
3. On treatment follow-up may be short in real-world practice, there is potential for informative censoring
    a. We have incorporated censoring weights
4. Some proposed secondary analyses may lack sufficient sample size
    a. Proposed a host of secondary and sensitivity analyses and in order of priority

9. Protection of human subjects

This study is part of a larger project called Leveraging real-world dAta to optimize PharmacotheRapy outcomes in multimorbid patients by using machine learning and knowledge representation methods (LEAPfROG project). The LEAPfROG project was exempted from requiring ethics approval (waiver W22_340 # 22.412) on September 2022, 22 by the Medical Ethics Committee of the Amsterdam University Medical Centres, University of Amsterdam, Amsterdam, the Netherlands, as it did not fall within the scope of the Dutch Medical Research Involving Human Subjects Act.

10. Reporting of adverse events

The proposed study is observational research that makes secondary use of data collected as part of routine care and does not involve any intervention or alteration in clinical care. Therefore, reporting of adverse events related to this study is not applicable. Safety evaluations for this study are limited to the specified safety outcomes stated in section 4.4.2.

11. References

Acute Boekje 2017-2025. Een initiatief van de Nederlandse Internisten Vereniging. Figuur Indeling en lokalisatie acute nierinsufficiëntie hetacuteboekje.nl/gerelateerde_documenten/bijlage/000411/1/1/Figuur Indeling en lokalisatie acute nierinsufficiëntie.html. Accessed 26 August 2025

Brinkman DJ, Hendriksen LC, Rigter IM, Hollmann MW. Pharmacology and relevant drug interactions of metamizole. *Br J Clin Pharmacol*. 2025; 1-8. doi:10.1002/bcp.70101

Chawla LS, Bellomo R, Bihorac A, Goldstein SL, Siew ED, Bagshaw SM, Bittleman D, Cruz D, Endre Z, Fitzgerald RL, Forni L, Kane-Gill SL, Hoste E, Koyner J, Liu KD, Macedo E, Mehta R, Murray P, Nadim M, Ostermann M, Palevsky PM, Pannu N, Rosner M, Wald R, Zarbock A, Ronco C, Kellum JA; Acute Disease Quality Initiative Workgroup 16.. Acute kidney disease and renal recovery: consensus report of the Acute Disease Quality Initiative (ADQI) 16 Workgroup. Nat Rev Nephrol. 2017 Apr;13(4):241-257. doi: 10.1038/nrneph.2017.2

Drewes, A.M., Jensen, R.D., Nielsen, L.M., Droney, J., Christrup, L.L., Arendt-Nielsen, L., Riley, J. and Dahan, A. (2013), Differences between opioids: pharmacological, experimental, clinical and economical perspectives. British Journal of Clinical Pharmacology, 75: 60-78. https://doi.org/10.1111/j.1365-2125.2012.04317.x

Fernández-Llaneza D, Hilbrands L B, Vogt L, Olde Engberink R H G, Klopotowska J E, Identifying a cohort of hospitalized chronic kidney disease patients using electronic health records: lessons learnt and implications for future research and clinical practice guidelines, *Clinical Kidney Journal*, Volume 18, Issue 4, April 2025, sfaf073, https://doi.org/10.1093/ckj/sfaf073

Fernández-Llaneza D, Vos RMP, Lieverse JE, Gosselt HR, Kane-Gill SL, van Gelder T, Klopotowska JE; LEAPfROG Consortium. An Integrated Approach for Representing Knowledge on the Potential of Drugs to Cause Acute Kidney Injury. Drug Saf. 2025 Jan;48(1):43-58. doi: 10.1007/s40264-024-01474-w

James MT, Levey AS, Tonelli M, et al. Incidence and Prognosis of Acute Kidney Diseases and Disorders Using an Integrated Approach to Laboratory Measurements in a Universal Health Care System. *JAMA Netw Open.* 2019;2(4):e191795. doi:10.1001/jamanetworkopen.2019.1795

Jolley RJ, Quan H, Jetté N, Sawka K. J, Diep L, Goliath J., Roberts D. J, Yipp B. G, Doig C. J. Validation and optimisation of an ICD-10-coded case definition for sepsis using administrative health data. *BMJ Open* 2015;**5:**e009487. doi: 10.1136/bmjopen-2015-009487

Joyce EL, DeAlmeida DR, Fuhrman DY, Priyanka P, Kellum JA. eResearch in acute kidney injury: a primer for electronic health record research. Nephrol Dial Transplant. 2019 Mar 1;34(3):401-407. doi: 10.1093/ndt/gfy052

Khwaja A. KDIGO clinical practice guidelines for acute kidney injury. Nephron Clin Pract. 2012;120(4):c179-84. doi: 10.1159/000339789

KNMP Kennisbank. 2025. https://kennisbank.knmp.nl/article/Informatorium_Medicamentorum/G806.html#G38. Accessed 2 June 2025.

Kuang A, Xu C, Southern DA, Sandhu N, Quan H. Validated administrative data based ICD-10 algorithms for chronic conditions: A systematic review. J Epidemiol Popul Health. 2024 Aug;72(4):202744. doi: 10.1016/j.jeph.2024.202744

Levey AS. Defining AKD: The Spectrum of AKI, AKD, and CKD. Nephron. 2022;146(3):302-305. doi: 10.1159/000516647

Mason SA, Nathens AB, Byrne JP, Fowler R, Gonzalez A, Karanicolas PJ, Moineddin R, Jeschke MG. The accuracy of burn diagnosis codes in health administrative data: A validation study. Burns. 2017 Mar;43(2):258-264. doi: 10.1016/j.burns.2016.11.008

McCaffrey, D.F., Griffin, B.A., Almirall, D., Slaughter, M.E., Ramchand, R. and Burgette, L.F. (2013), A tutorial on propensity score estimation for multiple treatments using generalized boosted models. Statist. Med., 32: 3388-3414. https://doi.org/10.1002/sim.5753

Stueber T, Buessecker L, Leffler A, Gillmann HJ. The use of dipyrone in the ICU is associated with acute kidney injury: A retrospective cohort analysis. Eur J Anaesthesiol. 2017 Oct;34(10):673-680. doi: 10.1097/EJA.0000000000000627

Susantitaphong, P, Cruz, D N., Cerda, J, Abulfaraj, M, Alqahtani, F, Koulouridis, I, Jaber, B. World Incidence of AKI: A Meta-Analysis. Clinical Journal of the American Society of Nephrology 8(9):p 1482-1493, September 2013. doi: 10.2215/CJN.00710113

Patrick AR, Schneeweiss S, Brookhart MA, et al. The implications of propensity score variable selection strategies in pharmacoepidemiology: an empirical illustration. Pharmacoepidemiol Drug Saf. Jun 2011;20(6):551-9. doi:10.1002/pds.2098

Pirracchio R, Petersen ML, van der Laan M. Improving propensity score estimators' robustness to model misspecification using super learner. Am J Epidemiol. 2015 Jan 15;181(2):108-19. doi: 10.1093/aje/kwu253

Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005 Nov;43(11):1130-9. doi: 10.1097/01.mlr.0000182534.19832.83

Quan H, Li B, Saunders LD, Parsons GA, Nilsson CI, Alibhai A, Ghali WA; IMECCHI Investigators. Assessing validity of ICD-9-CM and ICD-10 administrative data in recording clinical conditions in a unique dually coded database. Health Serv Res. 2008 Aug;43(4):1424-41. doi: 10.1111/j.1475-6773.2007.00822.x

Siew ED, Ikizler TA, Matheny ME, Shi Y, Schildcrout JS, Danciu I, Dwyer JP, Srichai M, Hung AM, Smith JP, Peterson JF. Estimating baseline kidney function in hospitalized patients with impaired kidney function. Clin J Am Soc Nephrol. 2012 May;7(5):712-9. doi: 10.2215/CJN.10821011

Turner, Neil N., and others (eds), *Oxford Textbook of Clinical Nephrology: Three-Volume Pack*, 4 edn (Oxford, 2015; online edn, Oxford Academic, 1 Oct. 2015), https://doi.org/10.1093/med/9780199592548.001.0001

Yasrebi-de Kom Izak A R, Dongelmans Dave A, Abu-Hanna Ameen, Schut Martijn C, de Lange Dylan W, van Roon Eric N, de Jonge Evert, Bouman Catherine S C, de Keizer Nicolette F, Jager Kitty J, Klopotowska Joanna E, Acute kidney injury associated with nephrotoxic drugs in critically ill patients: a multicenter cohort study using electronic health record data, *Clinical Kidney Journal*, Volume 16, Issue 12, December 2023, Pages 2549–2558, https://doi.org/10.1093/ckj/sfad160

Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA. 1998 Nov 18;280(19):1690-1. doi: 10.1001/jama.280.19.1690.

12. Appendices

See excel files.

Appendix A – Study population entry criteria (exposure)

Appendix B – Administration status

Appendix C – Administration routes

Appendix D – Drug-, diagnosis-, and procedure-based Inclusion and Exclusion Criteria

Appendix E – Drug, diagnosis-, and procedure-based covariates

Appendix F – Outcome