

---

# Hva er en genomvid assosiasjonsstudie?

---

## MEDISIN OG TALL

MARIE KLEVJER

marie.klevjer@ntnu.no

Marie Klevjer er ph.d. og forsker i medisin og helsevitenskap ved Cardiac Exercise Research Group (CERG) ved Institutt for sirkulasjon og bildediagnostikk, NTNU.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

STIAN LYDERSEN

Stian Lydersen er dr.ing. og professor i medisinsk statistikk ved Regionalt kunnskapssenter for barn og unge – psykisk helse og barnevern (RKBU Midt-Norge) ved Institutt for psykisk helse, NTNU.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

GURO F. GISKEØDEGÅRD

Guro F. Giskeødegård er professor i biostatistikk ved HUNT-senter for molekylær og klinisk epidemiologi ved Institutt for samfunnsmedisin og sykepleie, NTNU.

Forfatteren har fylt ut ICMJE-skjemaet og oppgir ingen interessekonflikter.

---

**I en genomvid assosiasjonsstudie (genome-wide association study, GWAS) forsøker man å identifisere genetiske varianter som er assosiert med en sykdom eller en fenotype. Vanligvis tester man flere millioner genetiske varianter fra hele genomet i repeterte regresjonsmodeller.**

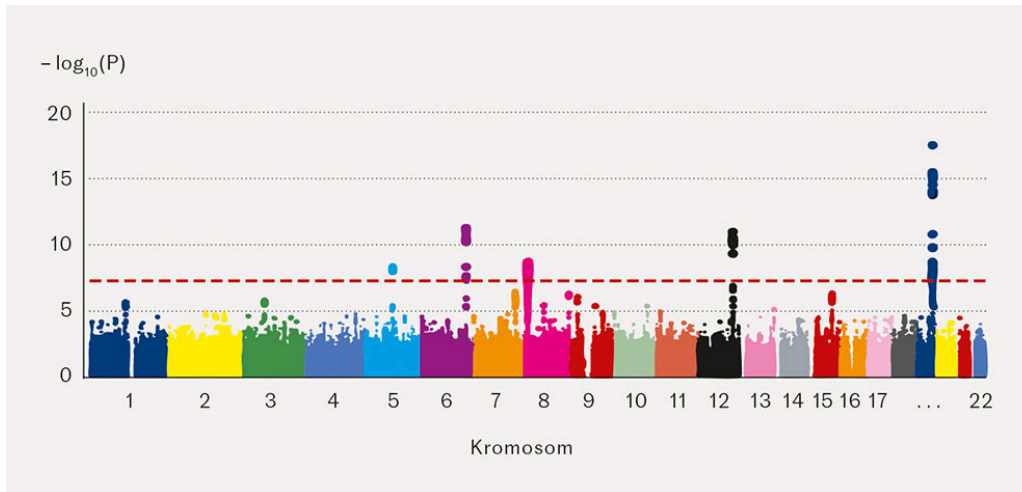
En genetisk variant refererer til en endring i nitrogenbasene (A, C, G, T) på en spesifikk posisjon i genomet. Forenklet kan man si at en genetisk variant forekommer dersom for eksempel flesteparten har en A på en spesifikk posisjon

i genomet, mens en mindre del av befolkningen har en T. I en genomvid assosiasjonsstudie ønsker man å undersøke om personer som har en T på denne plassen i genomet, har større eller mindre risiko for sykdom eller en fenotype (som f.eks. høyere kroppsmasseindeks) enn personer som har en A på den samme plassen. Dette gjøres for alle genetiske varianter som er tilgjengelige i datasettet, typisk flere titalls millioner.

---

## Manhattan-plott

I genomvide assosiasjonsstudier brukes et statistisk signifikansnivå på  $p < 5 \times 10^{-8}$ , og genetiske varianter som har en  $p$ -verdi mindre enn dette, regnes som statistisk signifikant assosiert med sykdommen eller fenotypen man undersøker. Man antar at det finnes ca. 1 million genetiske varianter som er uavhengig av hverandre (1). Ved å bruke signifikansnivå  $0,05 / (1\ 000\ 000) = 5 \times 10^{-8}$ , gjør man en Bonferroni-korreksjon for multiple tester, slik at sannsynligheten for å gjøre et falskt positivt funn holdes under 0,05 (2). Resultatet presenteres ofte i et Manhattan-plott, som viser alle genvariantene og tilhørende  $p$ -verdier, slik som i eksempelet i figur 1 (3).



**Figur 1** Et eksempel på et Manhattan-plott, hvor man har kromosomer og genetisk posisjon på x-aksen, og den negative logaritmen av  $p$ -verdien på y-aksen. Den røde stiplede linjen tilsvareer signifikansnivå på  $p < 5 \times 10^{-8}$ , og genetiske varianter over denne linjen regnes som signifikant assosiert med fenotypen. Figuren er modifisert fra Ikram og medforfattere (3). I offentlig eie

---

## Bruksområder

Resultater fra genomvide assosiasjonsstudier danner ofte grunnlaget for videre genetiske analyser. Resultatene brukes i hovedsak til å skaffe innsikt i de biologiske mekanismene til den aktuelle sykdommen eller fenotypen, men kan også brukes til å estimere arvelighet, regne ut genetisk korrelasjon – som er et mål på hvor mye det genetiske bidraget overlapper mellom ulike sykdommer eller fenotyper – samt identifisere kausale sammenhenger mellom risikofaktorer og sykdommer. Typiske videre analyser kan være funksjonelle

annoteringsanalyser, hvor man undersøker om de genetiske variantene påvirker gener, genuttrykk eller biologiske prosesser – for eksempel om de ligger i områder som regulerer genaktivitet, eller om de er knyttet til spesifikke celletyper eller vev. Man kan også gjøre mendelsk randomisering (4), hvor man kan bruke genetiske varianter fra en genomvid assosiasjonsstudie som instrument for å undersøke kausalitet, eller utvikle polygene risikoskårer (5), der man samler effekten av mange genetiske varianter som sier noe om den samlede genetiske tilbøyeligheten til sykdom eller fenotype. Dette kan igjen brukes som grunnlag for kliniske prediksjoner eller som bakgrunn for utvikling og testing av nye medisiner.

---

## Genetisk imputering

Ettersom kostnadene for genotyping har sunket i takt med den teknologiske utviklingen, har det blitt etablert biobanker med enorme mengder genetisk informasjon. I Norge er det stort sett store befolkningsstudier som brukes i genomvide assosiasjonsstudier. På verdensbasis er det UK Biobank som er det mest omfattende og mest brukte datasettet i sitt slag, bestående av 500 000 personer (6). De fleste befolkningsstudier har direkte genotypet noen hundre tusen genetiske varianter, mens resten – ofte flere titalls millioner – er imputert. Genetisk imputering er en metode som brukes for å *fylle inn* manglende genetisk informasjon ved å forutsi hvilke varianter som sannsynligvis finnes der vi ikke har direkte målinger, basert på mønstre i referansedata, noe som reduserer behovet for dyrere og mer omfattende analyser. Ulempen er at det kan forekomme feil i de estimerte variantene. Sjeldne varianter er ofte ikke godt representert i referansedataene, og genomvide assosiasjonsstudier er derfor lite egnet til å identifisere sjeldne mutasjoner.

---

## REFERENCES

1. Uffelmann E, Huang QQ, Munung NS et al. Genome-wide association studies. Nat Rev Methods Primers 2021; 1: 59. [CrossRef]
2. Lydersen S. Justering av p-verdier ved multiple hypoteser. Tidsskr Nor Legeforen 2021; 141. doi: 10.4045/tidsskr.21.0357. [PubMed][CrossRef]
3. Ikram MK, Sim X, Jensen RA et al. Four novel Loci (19q13, 6q24, 12q24, and 5q14) influence the microcirculation in vivo. PLoS Genet 2010; 6: e1001184. [PubMed][CrossRef]
4. Stensrud MJ. Mendelsk randomisering - genetisk tilnærming til epidemiologisk metode. Tidsskr Nor Legeforen 2016; 136: 1002–5. [PubMed][CrossRef]
5. Choi SW, Mak TSH, O'Reilly PF. Tutorial: a guide to performing polygenic risk score analyses. Nat Protoc 2020; 15: 2759–72. [PubMed][CrossRef]

6. Sudlow C, Gallacher J, Allen N et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med 2015; 12: e1001779. [PubMed][CrossRef]

---

Publisert: 28. august 2025. Tidsskr Nor Legeforen. DOI: 10.4045/tidsskr.25.0390  
Opphavsrett: © Tidsskriftet 2026 Lastet ned fra tidsskriftet.no 1. juli 2026.