דו״ח מדעי

מספר המחקר במשרד להגנת הסביבה:
161-2-9

שם המוסד המחלקה והמוסד המגישים:
הסתדרות מדיצינית הדסה

כותרת המחקר בעברית:
הקשר בין חשיפה בקרב מתבגרים לזיהום סביבתי במפרץ חיפה לבין מצב הבריאות בגיל 17 והארעות סרטן בבוגרים

כותרת המחקר באנגלית:
Association between adolescent exposure to environmental pollution in Haifa bay area, health status at age 17 and adult cancer incidence.

סוג הדו״ח (שנתי או מסכם): מסכם

מוגש ע״י

חוקרים ראשיים:
(בעמודה ימנית שם ותואר אקדמי, בעמודה השמאלית מוסד או מחלקה):

פרופ׳ ג׳רמי קרק ז״ל	הסתדרות מדיצינית הדסה
פרופ׳ רענן רז	האוניברסיטה העברית
ד״ר רונית זינרייך	הסתדרות מדיצינית הדסה
פרופ׳ גלעד טוויג	חיל הרפואה

חוקרים נוספים:
(בעמודה ימנית שם מלא ותואר אקדמי, בעמודה השמאלית מוסד או מחלקה):

פרופ׳ ליטל קינן-בוקר	המרכז הלאומי לבקרת מחלות
פרופ׳ דוד ברודאי	הטכניון

מוגש למדען הראשי
המשרד להגנת הסביבה

תודות: גב' אפרת יעקב, גב' רותי הררי-קרמר, מר זיו תלמי, חברת HERE, חברת סיסטמטיקס, פרופ' נועם לוין, גב' אסטלה דרזנה, ד"ר דורית צור, יובל, ד"ר רותי לב-בראור, פרופ' יעל דובובסקי, ד"ר אנדרי לוברסקי, פרופ' לאה בנטור, ד"ר אלכס גיללס-הלל, גב' הדס מגן-מולכו, ד"ר איזבלה קרקיס, ד"ר ארנה מצנר, ד"ר אילן לוי, ד"ר מאסימו סטפוג'יה, פרופ' אורה פלטיאל, ד"ר ירוסלב יוסים.

תקציר

רקע:

אזור מפרץ חיפה הוא אחד המטרופולינים הגדולים בישראל ומכיל אזור תעשייה גדול ומגוון הכולל בין השאר בית זיקוק ומתחם פטרו-כימי, תחנת כח ומפעלים רבים. מפרץ חיפה אינו חריג כיום בישראל מבחינת ריכוזים של מזהמי קריטריון באוויר (NO_x, SO_2, PM_{10}, $PM_{2.5}$, CO, O_3), אולם מרשם הפליטות לסביבה ונתונים אחרים מראים שמתקנים רבים במפרץ פולטים כמויות ניכרות של תרכובות אורגניות נדיפות (VOCs) ומתכות כבדות, שחלקן מזיקות לבריאות האדם. חומרים אלו אינם מנוטרים באופן מספק לצורך מחקר אפידמיולוגי, ולא נוטרו כלל בעבר. כמו כן, מרשם הפליטות איננו כולל מידע על החומרים הספציפיים הנפלטים, ורוב החומרים האורגנים הנדיפים מדווחים תחת קטגוריה כללית של VOCs שאינן מתאן (NM-VOCs). דבר זה אינו מאפשר להעריך אילו חומרים נפלטים ובאילו כמויות, אילו ריאקציות הם עוברים לאחר פליטתם לאטמוספירה, מהן החשיפות הכימיות הסביבתיות של האוכלוסיה כתוצאה מפליטות המתחם התעשייתי לאוויר וכיצד הן השתנו במהלך השנים. כתוצאה מכך, לא ניתן לבצע מחקר אפידמיולוגי קלאסי הכולל הערכת חשיפה אינדיווידואלית למזהמים ספציפיים במפרץ חיפה.

הספרות האפידמיולוגית והטוקסיקולוגית מעידה כי מגורים בקרבת מתחמים פטרוכימיים או חשיפה לחומרים אורגניים נדיפים מסוימים עלולות להגביר את הסיכון לאסטמה ולכמה סוגים של סרטן, בעיקר לוקמיה וסרטן הריאה. ניטורים אפידמיולוגיים דיווחו על היארעות עודפת של אשפוזים עקב אסתמה בילדים ושל מחלות סרטן בכלל האוכלוסייה במפרץ חיפה בהשוואה לאזורים אחרים בישראל. עם זאת, מחקרים אפידמיולוגיים קודמים במפרץ לא כללו הערכת חשיפה לזיהום התעשייתי, או שנעשו במערך מחקר אקולוגי הפגיע מאוד לבלבול (confounding) והטיות. מטרת המחקר הנוכחי הייתה לבחון קשרים בין חשיפה לזיהום אוויר תעשייתי במפרץ חיפה בילדות ובגיל ההתבגרות לתוצאים בריאותיים בגיל 17 ולסרטן במבוגרים.

שיטות:

זהו מחקר משולב. ראשית, ביצענו מחקר חתך לחקר קשרים בין זיהום אוויר תעשייתי במפרץ חיפה ומצב בריאותי בגיל 17. אוכלוסיית המחקר עבור מחקר החתך כללה את כל המתבגרים בגילאי 16-18 שנולדו בישראל ומצבם הרפואי הוערך לצורך גיוס צבאי על ידי חיל הרפואה הישראלי מ-1967 עד 2017 (N = 2,523,745), מהם 59% גברים. שנית, ביצענו מחקר עוקבה היסטורית כדי לחקור את הקשרים של זיהום אוויר תעשייתי במפרץ חיפה לסרטן במבוגרים. אוכלוסיית המחקר למחקר העוקבה כללה את כל הנבדקים ילידי ישראל שמצבם הרפואי הוערך לפני הגיוס לצבא על ידי חיל הרפואה בשנים 1967 עד 2012, בני 16-20 שנים בזמן הבדיקה הרפואית ולא אובחנו בסרטן בגיל הבדיקה (N = 2,187,317). זמן המעקב החל בתאריך הבדיקה והסתיים במועד אבחון הסרטן הראשון, מוות או סיום המעקב (31.12.2012), המוקדם מביניהם.

הערכת התוצא בגיל 17 התבססה על ההערכה הרפואית במרכזי הגיוס הצבאיים. עבור מחקר זה, השתמשנו באבחנות של אסתמה, נזלת אלרגית, דלקת עור אטופית ו-rhinoconjunctivitis (נזלת עם דלקת בלחמית העין),

ומדדים של אינדקס מסת גוף (BMI) ולחץ דם. כדי לזהות היארעות סרטן במבוגרים, קישרנו את נתוני היארעות הסרטן ממרשם הסרטן הלאומי בישראל ונתוני תמותה ממרשם התמותה הלאומי. בנוסף, חילקנו את כל אבחנות הסרטן ל-13 קטגוריות שונות.

כדי לייצג את השונות המרחבית בהתפלגות זיהום אוויר תעשייתי במפרץ חיפה על פני מפרץ חיפה במהלך העשורים האחרונים ולהקצות הערכת חשיפה לזיהום אוויר תעשייתי במפרץ חיפה, פיתחנו מודל המבוסס על אינטרפולציית קריגינג עם רזולוציה של 100 מ' וללא שונות טמפורלית, בהתבסס על נתונים של גופרית דו-חמצנית מתחנות הניטור באזור, בשנים מסוימות בהן ניתן היה לייחס ריכוזים אלו במידה רבה לפליטות ממקורות תעשייתיים (2002-2004).

עבור תוצאים בריאותיים בגיל 17, סיווגנו את החשיפה לארבע קבוצות שוות משתתפים במפרץ חיפה וקבוצת ייחוס אחת שכללה את כל הנבדקים משאר הארץ. פרופורציות ההימצאות ורווח הסמך (CI) של 95% חושבו לפי קטגוריות חשיפה עבור כל תוצא בריאותי, ומודלים ליניאריים כלליים המיישמים ניתוחי רגרסיה לוגיסטית רב-משתנית שימשו לבחינת הקשר המתוקנן בין זיהום אוויר תעשייתי במפרץ חיפה לאסטמה או מחלות אטופיות אחרות. כדי לבחון משתנים המשנים או מבלבלים/מערפלים את הקשר ביצענו גם מודלים מרובדים לפי עשור הלידה וארץ המוצא. כל המודלים תוקננו לשנת לידה, מצב סוציו-אקונומי ברמת האזור הסטטיסטי וזרם בית הספר, וחלק מהמודלים הותאמו בנוסף גם ל-BMI, סוג היישוב, תחמוצות חנקן (NOx) וחומר חלקיקי בעל קוטר הקטן מ- 2.5 מיקרומטר ($PM_{2.5}$).

עבור סרטן במבוגרים, הערכנו את הסיכון לסרטן באמצעות מודלים גולמיים ומתוקננים של סיכון פרופורציונלי של Cox. עבור ניתוחים אלה, השתמשנו בשלוש קטגוריות חשיפה בעלות מספר משתתפים שווים בתוך מפרץ חיפה (בנוסף לקטגוריית הייחוס מחוץ למפרץ). כמשתנים תלויים, השתמשנו תחילה במשתנה של סרטן כלשהו, ולאחר מכן הרצנו מודלים בסיסיים ומותאמים עבור כל קטגוריית סרטן בנפרד. קשרים דווחו כיחסי סיכון (HR) ו-95% CI עבור כל קטגוריית חשיפה לזיהום אוויר תעשייתי במפרץ חיפה בהשוואה לאוכלוסיה מחוץ למפרץ. תקננו את המודלים שלנו לקבוצת המשתנים הבאים: מין, שנת לידה, סוג יישוב, מוצא, תחמוצות חנקן וציון קוגניטיבי. מודלים מרובדים לפי תקופות לידה ומצב סוציו-אקונומי בדקו בלבול ואינטראקציות הקשורות לגורמים אלה. עוד ביצענו ניתוח רגישות שהגביל את המדגם המחקר למשתתפים עם בריאות תקינה בגיל 17.

תוצאות:

פרופורציית ההימצאות של אסתמה ונזלת אלרגית בגיל הבדיקה הייתה 5.91% ו-4.6%, בהתאמה, בעוד שההימצאות של מחלות אטופיות אחרות הייתה הרבה יותר נמוכה. פרופורציית ההימצאות של אסתמה בנבדקים שלא היו תושבי מפרץ חיפה הייתה 5.90%. פרופורציית ההימצאות של אסתמה הייתה גבוהה בהרבה במפרץ חיפה, אך הימצאות גבוהה זו התרכזה בשלוש הקטגוריות של חשיפה נמוכה יותר לזיהום אוויר תעשייתי במפרץ חיפה (6.8%, 6.9%, ו-6.9% בקטגוריות 1, 2, 3, בהתאמה), וקבוצת החשיפה הגבוהה ביותר לא סבלה מהימצאות עודפת של אסתמה (6.0%). הממצאים היו דומים במחלות אטופיות אחרות למעט אטופיק דרמטיטיס,

אשר עבורה לא נמצאה הימצאות עודפת באף אחת מקטגוריות החשיפה במפרץ, ואילו ההימצאות בקבוצת החשיפה הגבוהה ביותר הייתה נמוכה מההימצאות בקבוצת הייחוס של הנבדקים מחוץ למפרץ (0.40% לעומת 0.49%).
קשרים מתוקננים בין זיהום אוויר תעשייתי במפרץ חיפה לאסטמה או מחלות אטופיות אחרות הראו מגמה דומה של הימצאות מתוקננת עודפת, המוגבלת לקטגוריות החשיפה הנמוכות של זיהום אוויר תעשייתי במפרץ חיפה. תקנונים נוספים לא שינו את הצורה הכללית של עקומת חשיפה-תגובה, למעט תקנון ל- $PM_{2.5}$ שחיזק את כל מדדי הקשר בכל קטגוריות החשיפה במפרץ, אולם גם במודל זה הקשר החלש ביותר עם אסתמה נצפה בקטגוריית החשיפה הגבוהה ביותר. ההבדלים ב-BMI ויתר לחץ דם בין קבוצות החשיפה היו מינימליים ולא הציגו מגמה ברורה, והקשרים המתוקננים הראו תמונה דומה.
ריבוד אוכלוסיית המחקר על פי עשור הלידה הדגים שהקשר החלש ביותר עם אסתמה הוא עם קטגוריית החשיפה הגבוהה ביותר, ללא קשר לעשור הלידה (החל משנות ה-50 ועד לתחילת שנות ה-2000). עם זאת, הקשרים החזקים ביותר עם קטגוריות החשיפה האחרות במפרץ הודגמו בילידי שנות ה-90, ואילו הקשרים החלשים ביותר הודגמו בילידי שנות ה-70. ריבוד האוכלוסייה על פי שש קבוצות ארץ מוצא גילה עקומות חשיפה-תגובה בעלות צורה כללית דומה ללא קשר לארץ המוצא. גם חלוקה של החשיפה לשלוש קבוצות שוות גודל (בנוסף לקבוצת הייחוס) לא הדגימה קשרים שונים.
עוקבת הסרטן כללה מעקב של 41,696,278 שנות אדם. במשך 45 שנים, 47,129 משתתפים אובחנו כחולי סרטן, עם שיעור היארעות גולמי של 142.8 מקרים ל-100,000 שנות אדם בקבוצת הלא חשופים (תושבי שאר הארץ) ו-171.5, 171.7 ו-174.8 מקרים ל-100,000 שנות אדם בקבוצות החשיפה הנמוכה, בינונית וגבוהה, בהתאמה. הסרטן הנפוץ ביותר באוכלוסייה כולה, כמו גם באוכלוסיית מפרץ חיפה, היה סרטן שד, ולאחריו מלנומה. בקרב גברים, הסרטן הנפוץ ביותר היה באיברי הרבייה.
מודלים גולמיים של Cox הציגו קשרי חשיפה-תגובה חיוביים ומונוטוניים ברורים בין חשיפה לזיהום אוויר תעשייתי במפרץ חיפה לבין סרטן כלשהו, עם HR של 1.23 (95% CI: 1.17 - 1.29) עבור קטגוריית החשיפה הגבוהה ביותר, בהשוואה לקטגוריית הייחוס (תושבי שאר הארץ). הקשר נחלש מעט עם התקנון למבלבלים פוטנציאליים, עם HR של 1.16 (95% CI: 1.10 - 1.21) עבור קטגוריית החשיפה הגבוהה ביותר. הקשרים היו חזקים יותר עבור אלו שנולדו לאחר 1970 ובחמישון ה-SES התחתון.
בעת בחינת קשרים גולמיים של זיהום האוויר התעשייתי עם קטגוריות סרטן ספציפיות, המחלות הבאות הדגימו קשרים מובהקים סטטיסטית בדרגת החשיפה הגבוהה ביותר (לעומת קבוצת הייחוס של נבדקים משאר הארץ): מלנומה (HR = 1.52, 95% CI: 1.33 – 1.73), סרטן שד בנשים (HR = 1.20, 95% CI: 1.08 – 1.35), סרטן במערכת העצבים המרכזית (HR = 1.33, 95% CI: 1.07 – 1.64), סרטן בלוטת התריס (HR = 1.39, 95% CI: 1.16 – 1.66), ולוקמיה (HR = 1.36, 95% CI: 1.07 – 1.73).
התקנון של המודלים למשתנים מבלבלים/מערפלים כמעט ולא שינה דבר בתוצאות. הקשרים המתוקננים הדגימו עקומות חשיפה-תגובה חיוביות מונוטונית או כמעט מונוטונית עם קשרים שנחלשו מעט לעומת הקשרים הגולמיים, עבור 6 מתוך 13 קטגוריות של מחלות סרטן: סרטן השד בנשים, סרטן מערכת העצבים המרכזית, סרטן ראש

וצוואר, לוקמיה, מלנומה וסרטן בלוטת התריס. קטגוריות הסרטן הבאות לא היו קשורות באופן עקבי לחשיפה לזיהום אוויר תעשייתי במפרץ חיפה: סרטן באיברי מערכת העיכול, לימפומה של הודג'קין, לימפומה שאיננה הודג'קין, סרטן באיברי הרבייה (בגברים ובנשים), סרטן בדרכי השתן וסרטן ריאה. הקטגוריה היחידה עבורה נצפה סיכון מופחת באוכלוסיית שהתגוררה באזורי עם דרגת החשיפה הגבוהה ביותר הינה סרטן בדרכי השתן, אולם קשר זה לא היה מובהק סטטיסטית (HR = 0.90, 95% CI: 0.70 – 1.16). הגבלת המדגם לנבדקים עם בריאות תקינה בגיל ההתבגרות הראתה ממצאים דומים.

דיון:

הממצאים שלנו לגבי ההימצאות העודפת של אסטמה ומחלות אטופיות אחרות במפרץ חיפה מאששים דיווחים קודמים, אך הבדיקה המפורטת של הקשר בין חשיפה לזיהום אוויר תעשייתי במפרץ חיפה לבין מצבים רפואיים אלו אינה מעידה שעודף זה נגרם בגלל זיהום אוויר תעשייתי. עם זאת, ישנם מספר הסברים לממצאים אלו המתאימים לקשר סיבתי בין החשיפה לזיהום תעשייתי לבין אסתמה ומחלות אטופיות וניתן לבחון אותם במחקרים עתידיים. הסבר אפשרי ראשון הינו שממצאים אלו עשויים לנבוע מבלבול שיורי (residual confounding), למשל על ידי עישון, גורמים גנטיים או אחרים. הסבר שני הוא שהקשרים שהודגמו יכולים להיווצר כתוצאה מאינטראקציה של זיהום אוויר תעשייתי במפרץ חיפה עם גורמים שלא נמדדו במחקר, כגון חשיפה לאלרגנים. לדוגמה, במידה וזיהום האוויר התעשייתי במפרץ מגביר את הימצאות תחלואת האסתמה רק בנוכחות של חשיפה משמעותית לאלרגנים סביבתיים מסוימים, ואלרגנים אלו נדירים יותר בקבוצת החשיפה הגבוהה ביותר אך נפוצים באזורי החשיפה הנמוכה יותר במפרץ – יכולה להיווצר תמונת תחלואה דומה לזו שהודגמה במחקרנו.

הסבר אפשרי שלישי הוא נוכחות של הטיית מידע דיפרנציאלית הנובעת מביסוס הערכות החשיפה שלנו אך ורק על כתובות סביב גיל 16. מצב זה יכול להיווצר למשל אם משפחות עם ילדים אסתמטיים שחיו באזור החשיפה הגבוהה ביותר לזיהום אוויר תעשייתי במפרץ חיפה במהלך האבחון של ילדם עברו להתגורר באזורים אחרים לפני גיל 16, ותופעה זו הייתה נפוצה פחות במשפחות של ילדים אסתמטיים מאזורים אחרים. בהתאם, מכיוון שהמחקר מוגבל לנתוני חשיפה ותחלואה סביב גיל 16-17, עלולה להתקבל תמונה של הימצאות מופחתת בקבוצת החשיפה הגבוהה ביותר. מחקרים עתידיים יוכלו לבחון אפשרות זו בעזרת נתוני כתובות מפורטים יותר ונתוני היארעות אסתמה לאורך הילדות.

הקשרים שמצאנו עם מחלות סרטן תואמים השפעה סיבתית של זיהום אוויר תעשייתי במפרץ חיפה על הסיכון למחלות אלו. חוסר הקשר לסרטן הריאות והערמונית עשוי להיות מוסבר בכך שמחלות אלו שכיחות יותר בגילאים מבוגרים, ואילו עיקר זמן המעקב שלנו כולל גילאים צעירים. מצד שני, הקשרים של לוקמיה וסרטן מערכת העצבים המרכזית עם זיהום אוויר תעשייתי במפרץ חיפה תואמים מחקרים אפידמיולוגיים קודמים של קרבה לקומפלקסים פטרוכימיים אחרים, כמו גם עם הספרות המדעית לגבי תפקידם של ממסים וחומרים אורגניים נדיפים באטיולוגיה של גידולים אלו. קיימות אפשרויות שונות לכך שממצאים אלו הינם תוצאה של הטיות שונות. אפשרויות אלו נידונו בפירוט בגוף הדו"ח ואנו סבורים שהסיכוי שהן מסבירות לחלוטין את התוצאות, במיוחד

לאור תבנית הקשרים הנצפית עם מחלות הסרטן השונות ועם הסיכוי לסרטן כלשהו, הוא נמוך.
הסיכון המוגבר לסרטן כלשהו עשוי להצביע על כך שחשיפה לזיהום אוויר תעשייתי במפרץ חיפה מכילה מספר חומרים כימיים הפועלים במנגנונים ביולוגיים שונים כדי להגביר את הסיכון למחלות סרטן שונות. המחקר הנוכחי לא יכל לזהות חומרים או מנגנונים אלו מכיוון שלא ניתן היה לבצע הערכת חשיפה היסטורית למזהמים ספציפיים.
ובכל זאת, הוא מספק ראיות תומכות חדשות להשערה שחשיפה סביבתית לזיהום אוויר תעשייתי במפרץ חיפה העלתה את הסיכון לסרטן. לראיות אלו השלכות משמעותיות על בריאות הציבור, שכן אזור התעשייה במפרץ חיפה ממוקם בלבו של אזור מאוכלס.

Abstract

Background:
The Haifa Bay Area (HBA) is one of the largest metropolises in Israel and contains the country's principal industrial area. HBA has a dense air quality monitoring network; however, these stations continuously monitor mostly criteria pollutants. The national mandatory reported emissions inventory shows that many HBA facilities emit substantial amounts of volatile organic compounds (VOCs) and heavy metals, some of which are harmful to human health. In addition, the epidemiologic and toxicologic literature suggests that residential proximity to petrochemical complexes may increase the risk of asthma and some forms of cancer. However, previous epidemiological studies in HBA lacked exposure assessment for industrial pollution, or used an ecological design that it highly vulnerable to confounding and bias. We aimed to examine associations of childhood and adolescence residential exposure to HBA industrial air pollution (HBA-IAP) with health outcomes at age 17 and adult-onset cancer.

Methods:
This is a multi-method study. First, we undertook a cross-sectional study to study associations between HBA-IAP and health status at age 17. The study population for this design included all adolescents aged 16-18 born in Israel and whose medical status was evaluated for military recruitment by the Israeli medical corps from 1967 to 2017 (N = 2,523,745). Second, we used a historical cohort design to study HBA-IAP and adult-onset cancer associations. The study population for this design included all Israeli-born subjects whose medical status was evaluated prior to military recruitment by the medical corps from 1967 to 2012, aged 16-20 years at the time of the medical examination and did not have a cancer diagnosis at the age of examination (N = 2,187,317). Follow-up time started on the examination date and ended at the date of the first cancer diagnosis, death, or end of follow-up (31.12.2012), whichever came first.
Outcome assessment at age 17 was based on the medical evaluation at the military recruitment centers. For this study, we used diagnoses of prevalent asthma, allergic rhinitis, atopic dermatitis, and rhinoconjunctivitis, and measures of body-mass index and blood pressure. To assess adult-onset cancer, we linked cancer incidence data from the Israel National Cancer Registry and mortality data from the National Death Registry. We have further divided cancer diagnoses into 13 distinct categories. To capture the spatial variability in HBA-IAP distribution over the HBA area during the last decades and assign exposure assessment to HBA-IAP, we developed a kriging model with a 100 m resolution and no temporal variability, based on the mean 2002-2004 SO_2 concentrations in HBA air monitoring stations.
For health outcomes at age 17 we categorized the exposure to four equal-sized HBA groups and one reference group of participants outside the HBA. The crude prevalence and its 95% confidence interval (CI) were calculated by exposure categories for each health outcome, and generalized linear models applying multivariable logistic regression analyses were used to

examine the adjusted association between HBA-IAP and asthma or other atopic diseases. To examine effect modification and confounding possibilities further, we also performed stratified models by decade of birth and country of origin. All models were adjusted for year of birth, area-level socio-economic status, and school stream/orientation, and some models were further adjusted for BMI, type of locality, nitrogen oxides and $PM_{2.5}$.
For adult-onset cancer, we assessed the risk of cancer using crude and multivariable Cox proportional hazard models. For these analyses, we used the three exposure categories with equal numbers of participants within HBA (in addition to the non-HBA reference category). As dependent variables, we first used a variable of cancer of any type ("any cancer"), and then ran basic crude and adjusted models for each cancer category separately. Associations were reported as hazard ratios (HR) and 95% CI for each HBA-IAP exposure category compared with the non-HBA population. we adjusted our models to the following set of covariates: sex, year of birth, locality type, origin, NOx, and cognitive score. Models stratified by birth periods and socioeconomic status explored possible residual confounding and effect modifications by these factors. We further performed a sensitivity analyses limiting the study sample to participants with unimpaired health at age 17.

Results:
Asthma and allergic rhinitis prevalence were 5.9% and 4.6%, respectively, while other atopic diseases were much less prevalent. The prevalence of asthma was much higher in HBA, but this high prevalence was concentrated in the three categories of lower HBA-IAP exposure, and the residents in the highest HBA-IAP level did not show an excess of asthma prevalence. The findings were similar in other atopic diseases. Adjusted associations between HBA-IAP and asthma or other atopic diseases showed a similar trend of excess adjusted prevalence that is limited to the lower HBA-IAP exposure categories. Stratification by decade of birth or by country of origin, and further adjustments did not change the general shape of the curve. Differences in BMI and hypertension among the exposure groups were minimal and did not present a clear trend, and adjusted associations presented a similar picture.
The cancer cohort contributed 41,696,278 person-years of follow-up. Over 45 years, 47,129 participants were diagnosed with cancer, with a crude incidence rate of 142.8 cases per 100,000 person-years in the non-exposed category and 171.5, 171.7, and 174.8 cases per 100,000 person-years in the low, intermediate and high HBA-IAP exposure levels, respectively. In unadjusted Cox models, we observed clear positive exposure-response associations between HBA-IAP exposure and any cancer, with an HR of 1.23, 95% CI = 1.17 to 1.29 for the highest exposure category, compared to the reference category (non-HBA residents). The association slightly weakened when adjusting for potential confounders, with an HR of 1.16, 95% CI = 1.10 to 1.21 for the highest exposure category. The magnitude of the associations was stronger for those born after 1970 and in the lower SES quintile.
When examining specific cancer categories, adjusted models demonstrated positive monotonic or nearly-monotonic exposure-response curves for 6 out of 13 cancer groups: female breast cancer, CNS, head and neck, leukemia, melanoma, and thyroid cancer. On the other hand, the following cancer categories were not consistently associated with HBA-IAP

exposure: gastrointestinal organs, Hodgkin's lymphoma, NHL, reproduction organs (in males and females), urinary tract, and pulmonary cancer. Limiting the sample to those with unimpaired health in adolescence showed similar findings.

Discussion:
Our findings regarding the increased prevalence of asthma and other atopic diseases in HBA confirm previous reports, but our detailed examination of the relations between exposure to HBA-IAP and these medical conditions does not suggest that this increase is caused by air pollution from the industrial area. Nonetheless, the findings do not rule out causal relations between IAP in the Haifa Bay area and atopic diseases. These findings may result from residual confounding, for example by smoking or genetic factors. They can also be a result of an interaction of HBA-IAP with unmeasured factors such as exposure to allergens, if these factors are relatively absent in the highest exposure group. Another possibility is differential information bias that is a result of basing our exposure estimates solely on addresses aroung age 16. This could have happened if families with asthmatic children who lived in the highly exposed HBA-IAP areas during their child's diagnosis were more likely to change their address before age 16 than families of asthmatic children from other regions.
The associations with cancer are more consistent with a causal effect of HBA-IAP. The lack of association with pulmonary and prostate cancers may be explained because these diseases are more common in older ages. On the other hand, the associations of leukemia and CNS cancers with HBA-IAP are consistent with prior epidemiological investigations of proximity to petrochemical complexes other than HBA and with the literature regarding the role of solvents and VOCs in the etiology of these tumors. The increased adjusted risk of "any cancer" may suggest that HBA-IAP exposure contains several chemical agents that act in various biological mechanisms to increase the risk of various cancer diseases. Our study cannot detect these agents or mechanisms as it lacks exposure assessment for specific chemicals. Still, it provides new evidence for the hypothesis that residential exposure to HBA-IAP increased the risk of cancer. This evidence has major public health implications, as the industrial zone in HBA is located in a highly-populated area.

Keywords: Haifa Bay Area, volatile organic compounds, industrial air pollution, cancer, asthma, atopic diseases.

Introduction and Background

Haifa Bay Area

The Haifa Bay Area (HBA) is one of the largest metropolises in Israel and contains the country's principal industrial area. The latter includes both heavy industry (such as refineries and petrochemical complex, a power plant, petrol, and chemical storage and distribution centers) as well as a large number of small industrial plants (e.g., metal processing, chemical and pharmaceutical factories, garages, printing houses, etc.). Due to the development of the area, the historical physical separation between the urban areas, the agricultural fields, and the industrial zone (developed mainly in proximity to the seaport) has dramatically diminished in recent years. The HBA population is currently exposed to air pollutants from industrial, traffic, and agricultural sources. Furthermore, the unique topography of the area, with Mount Carmel in the southwest and the Zevolune hills in the northeast, results in complex local micro-meteorological conditions that affect the atmospheric distribution of the emitted pollutants.

Air quality in HBA

HBA has a dense air quality monitoring network; however, these stations continuously monitor mostly criteria pollutants (NO_x, SO_2, PM_{10}, $PM_{2.5}$, CO, and O_3). Based on the long record of these data, two important trends are noted: (1) a continuous reduction in concentrations of SO_2, and (to less extent) NO_x in the area, accompanied by an increase in O_3; (2) concentrations of the monitored criteria pollutants in HBA are generally lower than in other large cities in Israel, such as Tel Aviv and Jerusalem. Nevertheless, HBA contains a large industrial complex that emits other hazardous air pollutants. Examination of the national mandatory reported emissions inventory shows that many HBA facilities emit large quantities of volatile organic compounds (VOCs). For simplicity, VOCs and semi-VOCs (SVOCs) will be addressed hereafter as VOCs.

Most of the publicly reported VOC emissions are listed as "non-methane hydrocarbons" (NMHC or NMVOC), eliminating essential chemical information that is needed for estimating their environmental fate, environmental impact, and public health impact. The lack of detailed historical continuous monitoring of atmospheric VOCs is not surprising, considering that it may include hundreds of compounds at very low concentrations (ppt-ppb range) in such complex areas. Following this knowledge gap, the Israel Ministry of Environmental Protection has initialed in 2014 bi-weekly measurements of a several VOCs. The data available (with regard to both spatial distributions of sampling sites and to the number of compounds analyzed) leaves large uncertainty regarding VOCs distribution in the HBA and does not give a good picture of the population exposure to VOCs, certainly not for historical exposures that were not measured at all. Similarly, there is no detailed and wide monitoring of toxic metals in HBA, and the available measurements are mostly from recent years and are not always adequate with respect to their accuracy and sensitivity.

Health effects of VOCs and metals

In the long term, VOCs are mostly known for increasing the risk of various malignancies.[1,2] In the short term, and specifically for exposure through inhalation – which is the most relevant exposure pathway in the current context - VOCs such as aldehydes and polyaromatic hydrocarbons (PAHs) have been shown to irritate the respiratory tract in children,[2–5] and the mechanistic explanations for this effect are generally known.[4,6] Children are more sensitive to environmental pollution since their lungs are still developing, their airways are smaller, and they spend more time outdoors. A previous report showed that HBA presents the highest rate of pediatric hospitalizations due to asthma in Israel (114.5 per 100,000 children: more than twice the rate in the Tel Aviv area and 7-fold the rate in the Jerusalem area).[7] In addition, early childhood exposure to air pollution is also associated with an increased risk of respiratory morbidity later in life.[8]

Metals exposure may damage body lipids, proteins, and DNA through various mechanisms, such as the production of reactive oxygen and nitrogen species or binding to sulfhydryl groups and altering key cellular enzymes. Some of the metals (e.g., mercury, arsenic, cadmium, and lead) are systemic toxicants. They may cause clinical effects even in very low doses, while other metals serve as micronutrients that may be harmful only in high doses (e.g., copper, nickel, zinc, and cobalt). High levels of environmental exposure to specific metals were shown to have clinical effects on the nervous system, the kidneys, and the hematopoietic system, as well as carcinogenic and teratogenic effects, depending on the metal.[9] Metal contents in aerosols have also been associated with increased asthma symptoms and asthma medication use in several independent epidemiological studies,[10,11] a phenomenon supported by recent in-vitro investigations.[12]

Epidemiological studies of proximity to petrochemical industrial areas

The epidemiologic literature about associations between residential exposure to petrochemical industrial areas and health outcomes is sparse. Most of the epidemiological studies on this issue were either descriptive or ecological, limiting the ability to make firm causal conclusions. Still, most of these studies found an increased incidence or prevalence of various adverse health conditions with residential proximity to these areas. For example, such proximity was associated with an incident or prevalent asthma in studies from Taiwan,[13,14] South Africa,[15] and Argentina,[16] and with various other respiratory/atopic symptoms, such as wheezing, irritation, asthma exacerbations, allergic rhinitis, and respiratory hospitalizations.[14,16–21]

In addition to asthma and other atopic diseases, a major health concern with exposure to petrochemical industrial areas is cancer since petrochemical refineries emit VOCs, including some that are recognized as carcinogens. The main problems with most epidemiological investigations of this issue are the low number of cases of each specific cancer and the study design, which is usually ecological and almost always lacks detailed exposure assessment. Therefore, such studies are prone to residual confounding and typically suffer from low statistical power. Still, most studies on cancer and proximity to petrochemical industrial areas

found a higher incidence of either cancer or cancer mortality, with the leading associated cancer being leukemias[22–24] and respiratory (mainly lung) cancers.[22,25–27] Other studies found associations with liver[28] or pancreas[29] cancers, brain cancer,[30,31] bladder cancer,[27,30,32] and various hematological cancers.[27,33–35] In addition, studies that assessed association with any cancer (regardless of the cancer site) also found positive associations.[36–38] A comprehensive literature review of the epidemiological studies on associations between proximity to petrochemical industrial areas and adverse health outcomes, including outcomes that are beyond the scope of this study, can be found elsewhere.[39,40]

Epidemiological studies in HBA

Although measured and reported emissions show an improvement in air quality in HBA, there is substantial public concern about excess morbidity in HBA, in particular with respect to malignant, cardiovascular and respiratory diseases. These concerns are supported by the fact that HBA has had the highest age- and sex-adjusted cancer incidence in Israel since the late 1960s.[41,42] Given that these illnesses are affected by chronic exposures, current morbidity may have been impacted by past exposures that no longer exist in the area. Therefore, in order to study factors that influenced current, recent, or past morbidity, one has to obtain some exposure assessment based on historical measures.

Previous epidemiological studies mainly used various ecological study designs, in which both exposures and outcomes were assessed using aggregated measures without individual-level data.[43–46] Furthermore, the ecological studies examined associations with criteria pollutants (such as NO_x, SO_2, or PM) but not with VOCs or metals that were generally unmeasured (as discussed above). The ecological design in epidemiology is methodologically inferior since it suffers from the "ecological fallacy"[47] and is highly susceptible to confounding. As such, it cannot be used to test epidemiological hypotheses, and its primary use in contemporary epidemiology is limited to generating new hypotheses. In addition, criteria pollutants do not represent the unique exposures at HBA (as explained above) and cannot be attributed to the industrial zone in HBA.

One exception to the ecological design is a historical cohort study on cancer incidence in adults, in which individual-level cancer incidence data was used. The main limitation of that historical cohort study is that no exposure assessment was made, and risk was simply compared between HBA residents and non-HBA Israeli residents, with adjustment to several covariates but without considering the large HBA-IAP exposure variability that is expected among HBA resident.[48] Therefore, currently published environmental epidemiology studies from HBA cannot teach us much about the possible health effects of HBA specific emissions on the HBA population.

HBA: Open questions and major obstacles

Despite the high public interest in exposure to air pollution in HBA, many key questions on this issue remain open. For example, it is unknown how emissions from the HBA industry disperse in space and whether these exposures have any health effects on HBA residents.

These fundamental questions are highly dependent and require an innovative approach that will overcome several major obstacles:

1. Most VOCs and toxic metals are rarely measured, and the methods to measure them are technically complex and expensive.
2. The complex HBA topography makes it challenging to model spatial patterns of air pollution concentrations in the area. The cost of a full-scale, high spatial resolution model is prohibitive, and simple dispersion models without data assimilation of continuous monitoring data are insufficient.
3. Emissions from the HBA industry changed with time, and the early decades did not have any substantial air monitoring.

Study objectives

The main objective of this study is to examine associations between childhood and adolescence exposure to HBA-IAP with health status at age 17 and with adult-onset cancer. This study attempts to approach these questions using an innovative approach for exposure assessment: a new exposure model we developed to estimate historical exposure to air pollution from the HBA industry. In addition, we used the Israeli medical corps database (including health examinations during the years 1967 - 2017), Israel National Cancer Registry (INCR), and Israel National Death Registry data.

Methods

Study design and study population

This is a multi-method study. First, we used a cross-sectional design to study associations between HBA-IAP and health status at age 17. The study population for this design included all adolescents aged 16-18 born in Israel and whose medical status was evaluated for military recruitment by the Israeli medical corps from 1967 to 2017 (N = 2,523,745).
Second, we used a historical cohort design to study HBA-IAP and adult-onset cancer associations. The study population for this design included all Israeli-born subjects whose medical status was evaluated prior to military recruitment by the medical corps from 1967 to 2012, aged 16-20 years at the time of the medical examination and did not have a cancer diagnosis at the age of examination. Follow-up time started on the examination date and ended at the date of the first cancer diagnosis, death, or end of follow-up (31.12.2012), whichever came first. We excluded N=1560 adolescents with cancer diagnoses at the examination date or five years before that date, resulting in a study population of N=2,187,317 adolescents. Ethical permission for the study was granted by the institutional review board of the Israeli medical corps (protocol 1663-2011).

Outcome assessment

Identifications of health outcomes at regional military recruitment centers have been described previously.[49] Briefly, as part of the medical assessment for mandatory military service, adolescent males and females routinely undergo a medical evaluation approximately at the age of 17. The medical assessment includes a thorough review of medical history made by a physician and a physical examination that includes measurement of weight, height, and blood pressure. In general, abnormal findings from this evaluation are usually followed by additional tests or consultation, as appropriate. Subsequently, a specific standardized numerical code is assigned to each examinee to denote a medical diagnosis and is stored in a central database. This dataset was already used recently in several epidemiological studies.[49–54]

For health status at age 17, we focused most of our analyses on asthma, as this is a relatively prevalent condition with a substantial public health burden that is known to be influenced by air quality. We have also retrieved data about prevalent allergic rhinitis, atopic dermatitis, and rhinoconjunctivitis – various atopic conditions that share some mechanisms with asthma but also have other risk factors. Harmonization of these conditions was performed with medical corps experts, as they have changed over the years.
In addition, due to the growing literature on associations of air pollution with metabolic and cardiovascular outcomes, we used continuous data on body-mass index (BMI) and blood pressure as possible measures that may be affected by the pollution. BMI was categorized into four categories: underweight (<18.5 kg/m^2), normal (18.5-25 kg/m^2), Overweight (25-30 kg/m^2), and obese (>30 kg/m^2). Hypertension was defined as diastolic and systolic blood pressure of at least 90 and 140 mmHg, respectively.

To assess adult-onset cancer, we linked cancer incidence data from the Israel National Cancer Registry (INCR) up to 2012. The INCR meets internationally accepted requirements for the coding system (ICD-O-Version 3) and completeness of data, is considered to be ~ 97% complete for solid tumors, 88% complete for non-solid tumors, and has maintained consistently high coverage since its inception.[55] Clinical data include date and place of diagnosis, detailed tumor location (using the ICDO-3 codes), histopathology type, stage at the time of diagnosis, tumor size, lymph nodes involved, and information on treatments in the first six months after diagnosis. We linked mortality data from the National Death Registry, as stated above.

The classification of cancer diagnoses into meaningful clinical and epidemiological categories was defined by ICD-O topography codes, ICD-O-3 morphology codes, ICD-O-3 behavior codes, and statistical power considerations. The categories we decided to work with are detailed in Table 1 below.

Table 1. Cancer Site Groups and corresponding codes.

Category	**Sites**	**ICD-O *Topography* Codes**	**ICD-O-3 *Morphology* Codes.**	**ICD-O-3 *Behavior Codes***
Head and Neck	Lip Tongue Salivary Glands Gum & Hard Palate Floor of Mouth Buccal Mucosa Oropharynx Nasopharynx Hypopharynx Other Oral Cavity Nasal Cavities, Sinuses & Ear Larynx Eye (excluding melanoma)	C00.0-C00.9; C01.9-C02.9; C07.9, C08.0-C08.9; C03.0-C03.9, C05.0, C05.8,C05.9, C06.2; C04.0-C04.9; C06.0, C06.1, C06.8, C06.9, C11.0-C11.9, C12.9, C13.0-C13.9; C14.0-C14.8, C30.0-C30.9, C31.0-C31.9; C32.0-C32.9; C69.0-C69.9	Any valid code EXCEPT: Melanoma[1], Plasma cell tumors[2], Lymphomas [3], 9930	2,3
Gastrointestinal	Esophagus Stomach Small Intestine Colon Rectum/Anus Liver Gallbladder Pancreas (excluding endocrine) Other Digestive Tract	C15.0-C15.9; C16.0-C16.9; C17.0-C17.9; C18.0-C18.9; C19.9, C20.9, C21.0-C21.8' C22.0-C22.1, C23.9-C24.9, C25.0-C25.3, C25.5-C25.9	Any valid code EXCEPT: Melanoma, Plasma cell tumors, Lymphomas, 9930	2, 3
Pulmonary	Trachea Bronchus and Lung (Small and Non-Small Cell) Other Respiratory Sites	C33.9, C34.0-C34.9, C38.0-C38.8, C37.9, C39.0-C39.9	Any valid code EXCEPT: Melanoma, Plasma cell tumors, Lymphomas, 9930.	2, 3
Melanoma	Malignant Melanoma	C44.0-C44.9 or any other valid site, i.e., C51.0-C51.2, C60.0-C60.9, C69.0-C69.9, etc.	8720-8790	2, 3
Breast (female)	Breast (Female)	C50.0-C50.9. Female sex.	Any valid code EXCEPT: Melanoma, Plasma cell tumors,	2, 3

			Lymphomas, 9930.	
Reproduction (female)	Cervix (excluding CIN1/2) Endometrium (Corpus Uteri) Ovary Other Female Genital Organs	C53.0-C53.9; C54-C54.9; C56.9; C52.9, C55.9, C58.9, C57.0-C57.9, C51.0-C51.9	Any valid code EXCEPT: Melanoma, Plasma cell tumors, Lymphomas, 9930.	2, 3 Only 3 for C53.0-C53.9
Reproduction (male)	Prostate Testis Other Male Genital Organs	C61.9, C62.0-C62.9, C60.0-C60.9, C.63.0-C63.9	Any valid code EXCEPT: Melanoma, Plasma cell tumors, Lymphomas, 9930.	2, 3
Urinary	Bladder Kidney Other Urinary Organs	C67.0-C67.9; C64.9; C65.9, C66.9, C68.0-C68.9	Any valid code EXCEPT: Melanoma, Plasma cell tumors, Lymphomas, 9930.	2, 3
Central Nervous System (CNS)	Brain Other CNS	C71.0-C71.9;, C70.0-C70.9, C72.0-C72.9; C70.0-C72.9, C75.1-C75.3	Any valid code EXCEPT: Melanoma, Plasma cell tumors, Lymphomas, 9930. For C70.0-C72.9, C75.1-C75.3 – Any valid code	2, 3 For C70.0-C72.9, C75.1-C75.3: 0, 1.
Thyroid	Thyroid	C73.9	Any valid code EXCEPT: Melanoma, Plasma cell tumors, Lymphomas, 9930.	2, 3
Hodgkin Lymphoma	Hodgkin Lymphoma	C77.0-C77.9 or any valid extranodal site	9650-9667	3

Non-Hodgkin Lymphoma	Non-Hodgkin's Lymphomas Chronic Lymphocytic Leukemia Mycosis Fungoides	C77.0-C77.9 or any valid code	9590-9596, 9670-9699,9702-9719,9727-9729 and 9827 unless w/C42	3
Leukemia	Lymphoid Leukemias Myeloid Leukemias Other Leukemias Myeloproliferative Myelodysplastic Disease	C42.0-C42.4	9800-9827, 9831-9920; 9931-9948	3

[1] 8720-8790.
[2] 9731-9734.
[3] 9650-9667, 9590-9596, 9727-9729, 9670-9699, 9827 unless with C42, 9702-9719.

Exposure assessment

Geocoding

Our initial geocoding was performed by an external company (Taldor), which geocoded approximately 80% of the addresses up to the year 2011. We also merged the subjects with Central Bureau of Statistics (CBS) census tract SES measures using the resulting coordinates. In the second phase, we performed the geocoding process and the linkage to environmental exposures at the Hebrew University for the years 2012-2017. Overall, 79.8% of the addresses were geocoded at the full address or street level), and 10.0% were geocoded at the neighborhood or the locality level. Most of the remaining addresses were incomplete or contained errors that could not be corrected, and data for these adolescents were not analyzed. In HBA, we were able to geocode 89.1% of the addresses at the full address or street level. In addition, we have randomly sampled 1000 addresses geocoded by Taldor and geocoded them at HUJI to validate the geocoding results process for the main study population (cohorts up to 2011). From the addresses in the sample that were geocoded by Taldor (996/1000, 99.6%), we found 88.9% (885/996) of pairs of coordinates to be very close (< 1000m) to each other, and another 9.5% (95/996) of those pairs to be in the distance less than 5000m (additional details can be found in Table 2 below).

Table 2. Results of the geocoding validation process for the main sample.

		Locality Population Size			
Level of geocoding	**Distance difference (m)**	**< 10,000**	**10,000-10,0000**	**> 100,000**	**Total**
Full address	**1-1000**	41	284	351	676
	1000-5000	0	1	16	17
	5000-10000	0	0	2	2
	10000-50000	1	1	0	2
Street and locality	**1-1000**	17	39	24	80
	1000-5000	2	9	6	17
	5000-10000	0	0	3	3
	10000-50000	0	0	0	0
Locality center	**1-1000**	102	23	5	130
	1000-5000	13	23	25	61
	5000-10000	0	0	5	5
	10000-50000	1	0	0	1
	>50000	2	0	0	2
Total		179	380	437	996

HBA industrial air pollution

For this study, we have developed a model of exposure to HBA-IAP with a 100 m resolution and no temporal variability. The purpose of the model is to capture the spatial variability in HBA-IAP distribution over the HBA area during the last decades, given the limited actual data available about this issue (as described above). We will provide here a brief summary of the model.

Rationale: A comprehensive industrial emissions report has been published by the Ministry for Environmental Protection only since 2011. Some earlier emission data were collected by the Haifa Bay Municipality Association for the Environment, but they are not as detailed, and even they do not go back to the 20th century. It is thus impossible to contemplate the construction of a dispersion model for the HBA industrial emissions in earlier decades. Moreover, HBA's complex topography might not have enabled air pollution dispersion modeling with an acceptable accuracy at a reasonable cost. Using statistical models based on air pollution monitoring data is a common alternative, usually preferred, to modeling the spatial distribution of pollutants. Unfortunately, monitoring in HBA had not commenced before the beginning of the 1990s and the number of pollutants observed then was minimal. It is thus not practical to estimate spatial maps of an extensive range of contaminants from which actual concentration values could be extracted and assigned to the study subjects.

Given these limitations, we provided a rough scheme that will enable us to divide HBA into several categories of exposure to air pollution emitted from the HBA industrial area.
Monitoring data: During the development of the model, the air quality array in HBA included 26 monitoring stations observing a large number of air pollutants and meteorological variables. We used only the SO_2 concentrations data which were observed in a maximum of 21 stations. Figure 1 shows the locations of the SO_2 monitoring stations, and Figure 2 shows their temporal coverage since 1991. Wind data (direction and speed) were taken from the Israeli Meteorological Service (IMS) Afeq station, measured at 10-minute resolution from 2003-2016. To examine possible historical changes in the typical wind patterns, wind data at a three hours resolution were extracted for 1971-2000 from the Bet Dagan station (100 km south of HBA). This station was replaced in the early 2000s by an automatic station observing at a 10-minutes time resolution. Three hours resolution data were extracted from its records to continue the records of the historic station. Planetary boundary layer (PBL) characteristics, based on the Bet Dagan radiosonde data for 1970-2017, were obtained from the Integrated Global Radiosonde Archive (IGRA), maintained by the USA National Oceanic and Atmospheric Administration (NOAA). The radiosonde is launched twice a day at UST noon and midnight (14:00 and 02:00 local time, respectively), but the vertical resolution of the IGRA data is not sufficient for probing the night–time PBL and thus, only the afternoon records were used.

Assumptions and principles of the HBA-IAP exposure assessment model:

1. Exposure to pollution from HBA industrial emissions occurs only in locations to which the emitted pollution can be transported by air flow that previously passed by the industrial stacks.
2. Since the 1990s, heavy industry and the port have been the only emitters of SO_2 in HBA. In recent years, with the port the only significant source of SO_2 in HBA, the SO_2 levels have been very low and close to the instrument detection limit. Thus, it can be assumed that SO_2 observations from the beginning of the 21st century were dominated by heavy industry emissions and can serve as tracers for the industrial dispersion in the area.
3. Dispersion of industrial pollutants with an atmospheric lifetime similar to or longer than that of SO_2 is expected to follow a similar dispersion pattern to SO_2. Background SO_2 concentrations can be considered homogeneous in space. Thus, the spatial concentration pattern of pollutants with a long atmospheric lifetime emitted by the HBA industry is expected to roughly follow a similar spatial pattern to SO_2 concentrations.
4. Dispersion of industrial pollutants with an atmospheric lifetime shorter than that of SO_2 (e.g., most VOCs) may result in different spatial patterns. Still, it is expected that the principal axes of their dispersion lobes will be along those of the SO_2 spatial distribution.
5. Use of the observed SO_2 concentrations in the 21st century to assess the exposure to industrial pollution emitted in the last decades of the 20th century is warranted only if the relevant meteorological factors, mainly the wind patterns, have not significantly changed.
6. Interpolation concentrations maps rely on observed values. In spatially sparse observations, the high peaks and low troughs in an interpolation map reflect the

observations locations, not true concentration features. Thus, reliance on the small-scale details of such maps should be avoided, and only the general outline of the main patterns be used. (This is the main reason for categorizing our exposure assessment in our epidemiological models).

Spatial maps of the SO_2 concentrations should be produced using as many data points as possible to minimize the dependence of the map's patterns on the locations of each of the observation locations. Minimizing the reliance of the map features on the observation values can also be achieved by using an interpolation scheme that allows for observation errors. Additional consideration should be the SO_2 levels. With SO_2 levels in recent years becoming as low as the measuring instruments' detection limits, the specific annual calibration of the instruments might impact the results. Using data from years when the SO_2 values were higher is a way to minimize this effect. Given these considerations, the best trade-off which was found is the construction of an exposure map using kriging interpolation of the mean 2002-2004 SO_2 data. In this period, the number of observation locations was maximal (N = 21), with relatively few periods of missing data. It was a period when the SO_2 concentrations were still relatively high and for which optimized spatial pollution interpolation maps for SO_2 exist and can serve as a reference.[56]

Asaf et al. (2008) estimated the background levels of pollutants in Israel in three field campaigns during winter 2002, summer 2003, and spring 2005.[57] The background SO2 values which they report are 2.6, 1.4, and 1.1 ppb, respectively. Given that the campaigns were relatively short (three weeks each), they might reflect temporary values depending on the dominating meteorology during the campaigns. This is especially true of the winter and spring values observed during seasons with high meteorological variability. The summer season in Israel is very stable, so the value of 2.6 ppb can be regarded as more reliably representing the SO2 background concentration value for this season during the early 2000s. The period of 2002-2004 also coincides with the period of the Asaf et al. (2008) background pollution measurement campaigns.[57] Sensitivity of the SO_2 spatial patterns to the selected period of years was carried out. In all cases, the kriging nugget effect was set to 1 ppb to allow for the 1 ppb uncertainty in the observations reported by Yuval and Broday (2006).[56]

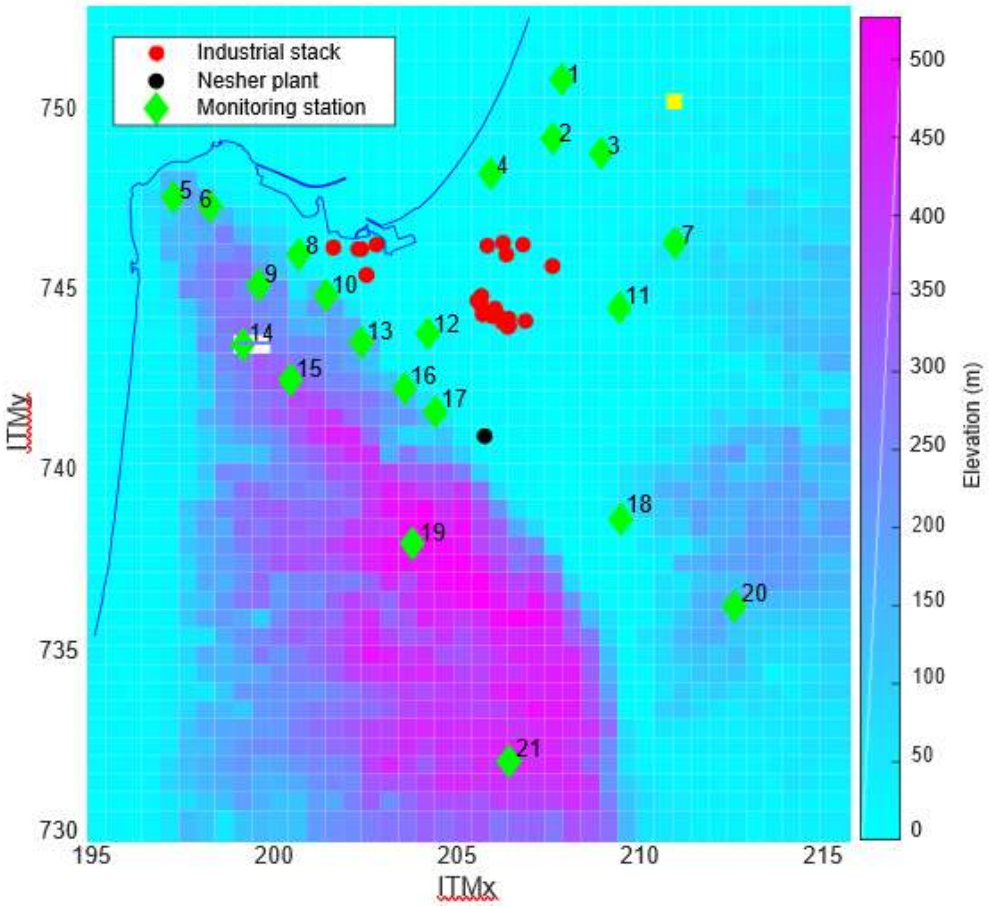

Figure 1. Study area map. Superimposed on the HBA elevation map are industrial emission stack locations from the 2015 Ministry of Environmental Protection report (red circles), the location of the Nesher cement plant furnace (black circle), the location of the Afeq meteorological station (yellow square) and the locations and numbers of the SO_2 monitoring stations whose data were used (green diamonds).

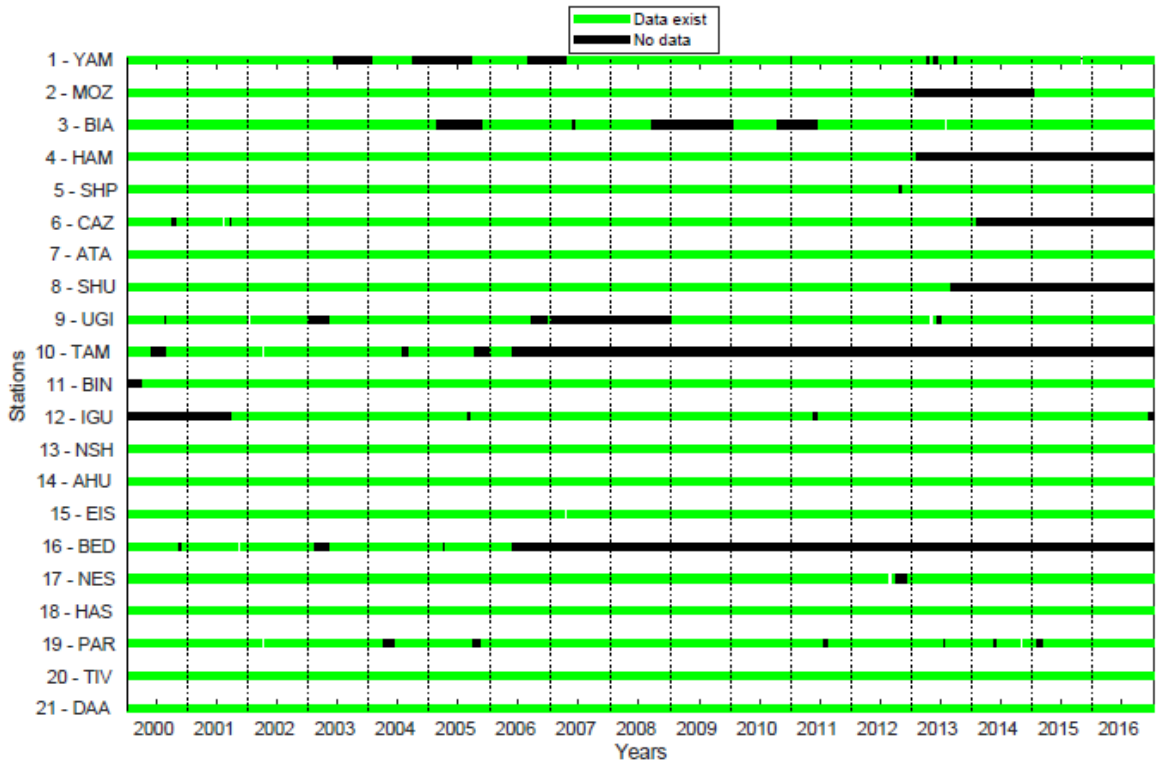

Figure 2. The temporal coverage of the SO_2 records observed by the stations in the study area since 2000.

Statistical analyses

Health outcomes at age 17

The crude prevalence and its 95% confidence interval (CI) were calculated by exposure categories for each of the health outcomes, and these measures were presented using simple bar plots or tables. The main categorization of the exposure was four equal-participants HBA groups and one reference group of participants outside HBA. Additional categorizations used in sensitivity analyses were three and four equal-interval exposure groups.
Generalized linear models applying multivariable logistic regression analyses were used to examine the adjusted association between HBA-IAP and asthma or other atopic diseases. Associations were expressed as odds ratios and 95% CI for each HBA-IAP exposure category compared to the non-HBA population, which we defined as the reference category. Models of BMI categories used multinomial (instead of logistic) regression.
All models were adjusted for the following set of potential confounders:

- **Year of birth**; Socio-economic status (**SES**, taken from the Central Bureau of Statistics at the level of small statistical area); and **school orientation** (Secular, religious, ultra-orthodox).

Some of the models were further adjusted for two criteria air pollutants:

- Nitrogen oxides (**NOx**, as a measure of traffic-related air pollution, taken from a land-use regression model by Dr. Ilan Levy).[58,59] This national model assesses annual averages of NOx levels over Israel since 1961, at a spatial resolution of 200 m over the entire domain and 50 m in urban regions.
- Particulate matter with diameter < 2.5 micrometer (**$PM_{2.5}$**, taken from a hybrid satellite-based model by Prof. Itai Kloog).[60,61] This national model assesses daily averages of $PM_{2.5}$ levels over Israel since 2003, at a spatial resolution of 1 km. Since the $PM_{2.5}$ model only starts from 2003, models that were further adjusted for $PM_{2.5}$ included only adolescents born from that year and on (approximately 35% of our study population).

To assess potential confounding by time trend, we performed stratified analyses of the sample by decades. This also enabled exploration of possible effect modification since the exposure model lacks any temporal variability (since the exposure model is time-invariant). An additional stratification by parental origin served us to assess residual confounding by genetic and cultural factors. A Jewish "Ashkenazi" origin was defined as birth in one of the following countries: Russia and other former USSR countries, Poland, Hungary, Romania, Germany; similarly, a Jewish "Mizrahi" origin was defined as birth in Morocco, Iraq, Egypt, Turkey, Yemen, or Iran. We also tested the effect of further adjustment for BMI and locality type (rural, urban with <50,000 residents, and urban with >50,000 residents).

Cancer

We assessed the risk of cancer using crude and multivariable Cox proportional hazard models. For these analyses, we used the three equal participants number exposure categories within HBA (in addition to the non-HBA reference category). With this decision, we increase the number of specific cancer cases in each HBA exposure category (which is the factor that limits statistical power in these analyses) while still being able to examine exposure-response relationships. As dependent variables, we first used a variable of cancer of any type ("any cancer"), for which the high number of cases permits various sub-group and sensitivity analyses. In addition, we ran basic crude and adjusted models for each cancer category separately. Associations were reported as hazard ratios (HR) and 95% confidence intervals (CI) for each HBA-IAP exposure category compared with the non-HBA population.
To control potential confounding, we adjusted our models to the following set of covariates: sex, year of birth, locality type, origin, NOx, and cognitive score (evaluated at examination date by the general intelligence test). Models stratified by birth periods and socioeconomic status explored possible residual confounding and effect modifications by these factors. To examine the possibility of residual confounding by coexisting illness, we performed sensitivity analyses limiting the study sample to participants with unimpaired health at the examination date, defined as lack of chronic treatment or need in medical follow-up and no history of cancer or major surgery.

All analyses were conducted in R, version 3.6.1, with the survival models fitted and tested for violations of proportionality assumption (by interaction terms with time) using the cox.zph function in the survival package.[62]

Results

Health outcomes at age 17

Study population

The study population included N = 2,523,745 adolescents who were born in Israel between the years 1947-2001, among which 59% were males. Asthma and allergic rhinitis prevalence were 5.9% and 4.6%, respectively, while other atopic diseases were much less prevalent (Table 3).

Table 3. Descriptive statistics of the study population (N = 2,523,745).

Characteristic	N (%)
Male sex	1492361 (59)
Year of birth	
1947-1949	11,283 (<1)
1950-1959	309,236 (12)
1960-1969	391,228 (15)
1970-1979	563,971 (22)
1980-1989	574,040 (23)
1990-1999	619,558 (24)
2000-2001	54,429 (2)
School orientation	
Secular	1,903,096 (75)
Religious	206,645 (8)
Ultra-orthodox	44,735 (<2)
Missing	369,269 (15)
Diagnosis	
Asthma	149,109 (5.9)
Atopic rhinitis	116,229 (4.6)
Rhinoconjunctivitis	10,126 (0.4)
Atopic dermatitis	12,506 (0.5)
Hypertension	
Yes	12,364 (0.5)
No	2,145,981 (85.0)
Missing	365,400 (14.5)
BMI (kg/m^2)	
Underweight (<18.5)	328,508(13.0)
Normal (18.5-25)	1,750,191(69.3)
Overweight (25-30)	279,283(11.1)
Obese (>30)	85,259 (3.4)
Missing	80,504(3.2)

Exposure levels

As explained above, participants were categorized into four equal-participants HBA-IAP exposure groups by their residential address. The number of participants in each group is specified in Table 4, and the spatial division of the area by these categories is shown in Figure 3.

Table 4. Main HBA-IAP exposure categories.

HBA-IAP category	N
Reference (non-HBA)	2,216,927
1	56,480
2	56,637
3	56,587
4	56,612

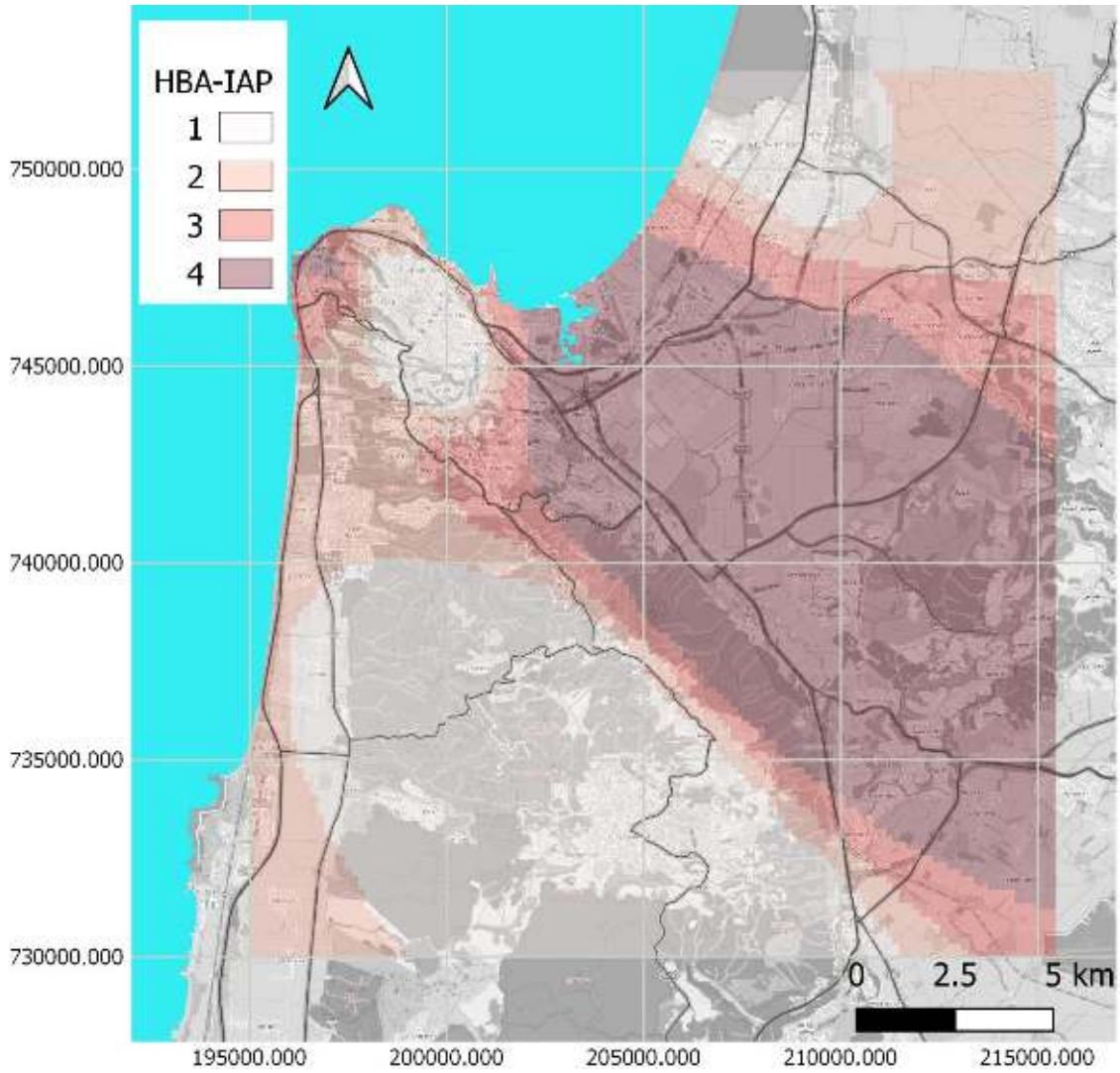

Figure 3. A map of the HBA-IAP exposure model, with color-coded exposure categories 1-4.

Crude prevalence of health conditions at age 17 by exposure categories

Asthma

To examine the burden of various health conditions by exposure and explore initial, crude associations, we first describe the prevalence of these health conditions and their 95% CI using simple bar plots. Figure 4 represents such a plot for asthma and shows that the prevalence of the disorder is indeed much higher in HBA, but this high prevalence is concentrated in the three categories of lower HBA-IAP exposure, and the residents who are assumed to be exposed to the highest levels of HBA-IAP do not represent excess asthma prevalence. The exact details of these analyses can be found in Supplementary Table S1.

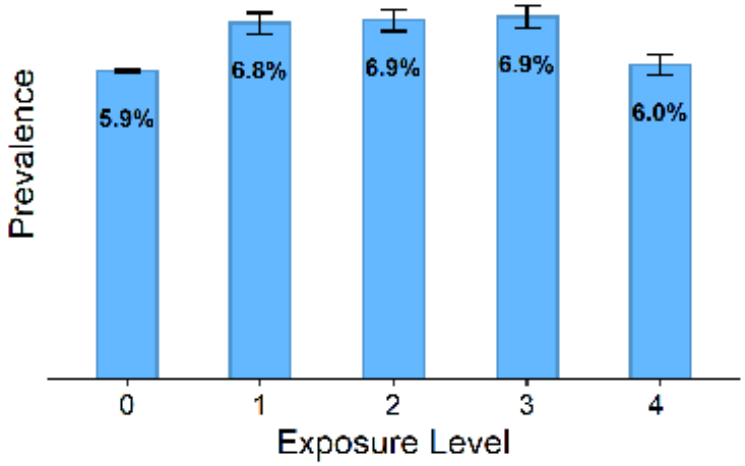

Figure 4. Crude prevalence of asthma by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

It is interesting to compare the shape of the curve in Figure 4 to that of a similar curve, when the exposure is taken as categories of known hazardous air pollutants. Figures 5 and 6 present such plots for $PM_{2.5}$ and asthma and for NOx and asthma in our study population, respectively and regardless of HBA-IAP. Both figures clearly demonstrate increasing prevalence by concentration. Exact details of these analyses are found in Supplementary Tables S2 and S3.

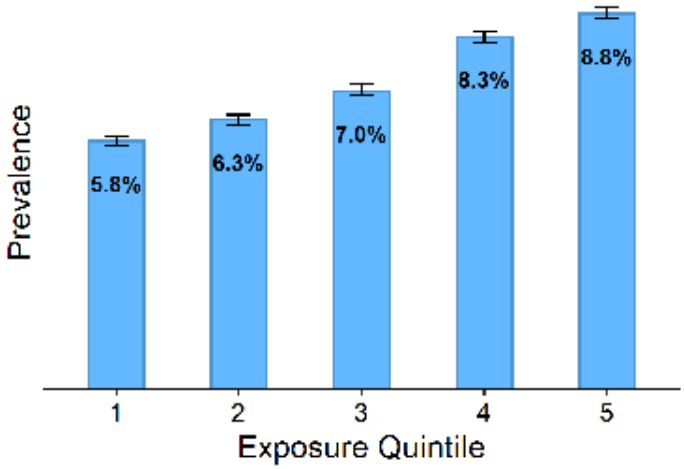

Figure 5. Crude prevalence of asthma by $PM_{2.5}$ quintiles for the entire study population.

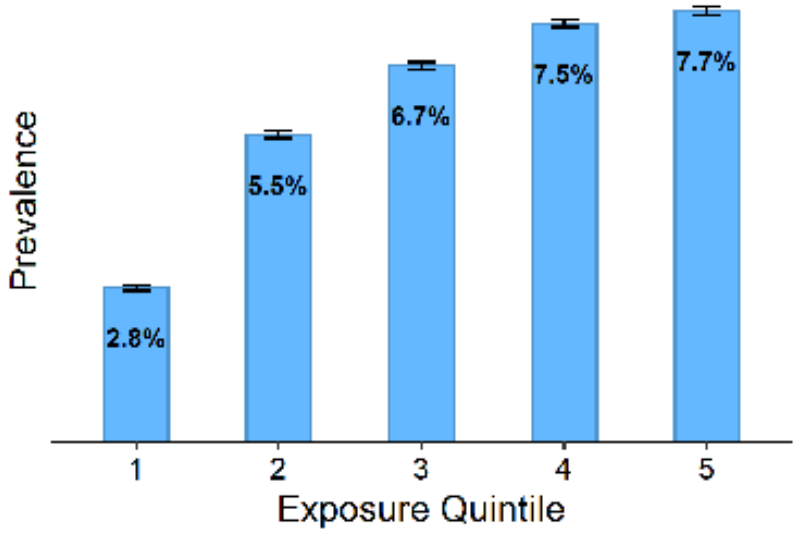

Figure 6. Crude prevalence of asthma by NOx quintiles for the entire study population.

Other atopic diseases

The prevalence of rhinitis, asthma with comorbid rhinitis, atopic dermatitis, and rhinoconjunctivitis by HBA-IAP categories is presented in Figures 7-10, respectively. HBA adolescent residents show an increased prevalence of rhinitis (with or without asthma) and rhinoconjunctivitis. However, similarly to asthma, the prevalence does not increase with increasing exposure to HBA-IAP, and the highest HBA-IAP exposure group presents the lowest prevalence level within HBA, and sometimes also in comparison to non-HBA residents. The exact details of these analyses can be found in Supplementary Table S4.

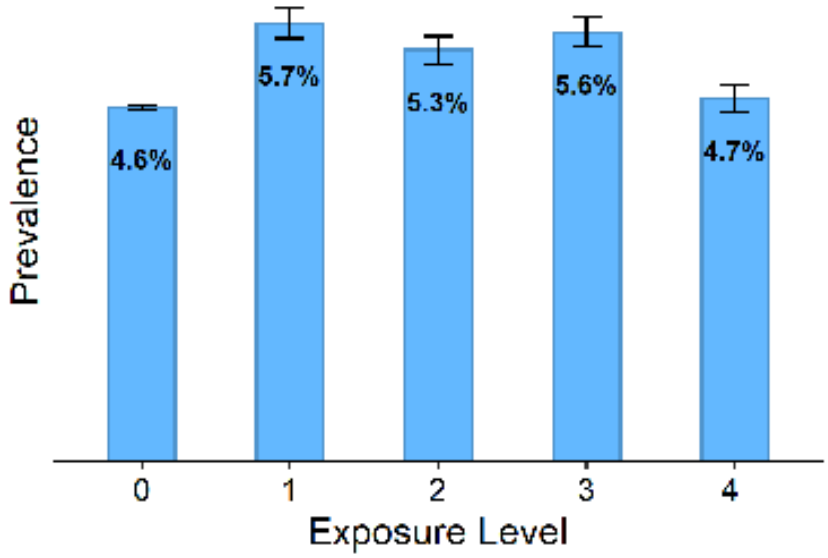

Figure 7. Crude prevalence of rhinitis by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

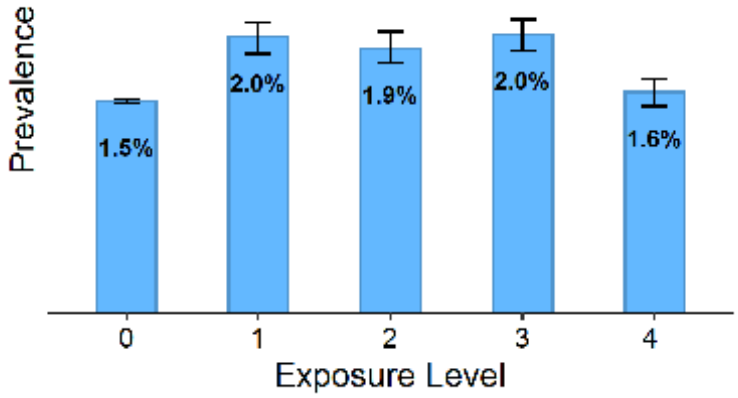

Figure 8. Crude prevalence of asthma with comorbid rhinitis by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

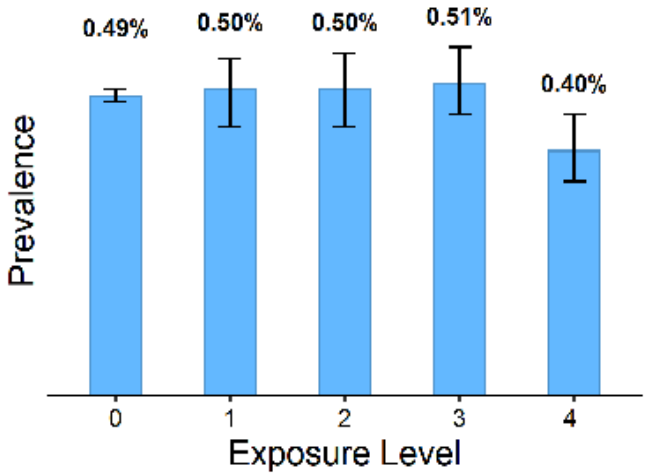

Figure 9. Crude prevalence of atopic dermatitis by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

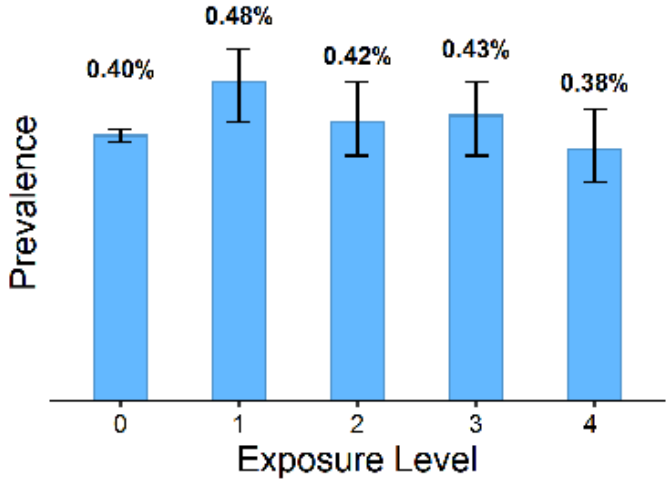

Figure 10. Crude prevalence of rhinoconjunctivitis by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Metabolic outcomes

The prevalence of underweight, overweight, and obesity by HBA-IAP categories is presented in Figures 11, 12, and 13, respectively. The exact details of these analyses can be found in Supplementary Table S5. Even though the highest exposure group presents a slightly elevated prevalence of obesity, the differences in BMI categories among the exposure groups are minimal. Likewise, differences in hypertension prevalence by exposure are minimal and do not present a clear trend (Figure 14). The exact details of these analyses can be found in Supplementary Table S6.

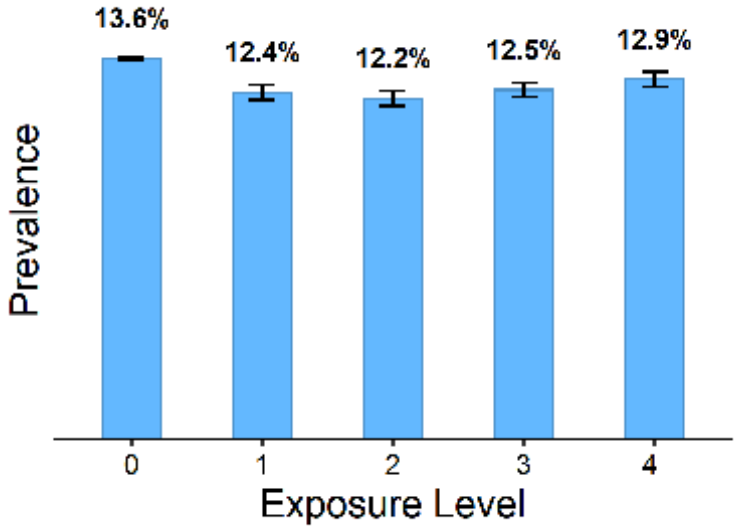

Figure 11. Crude prevalence of underweight by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

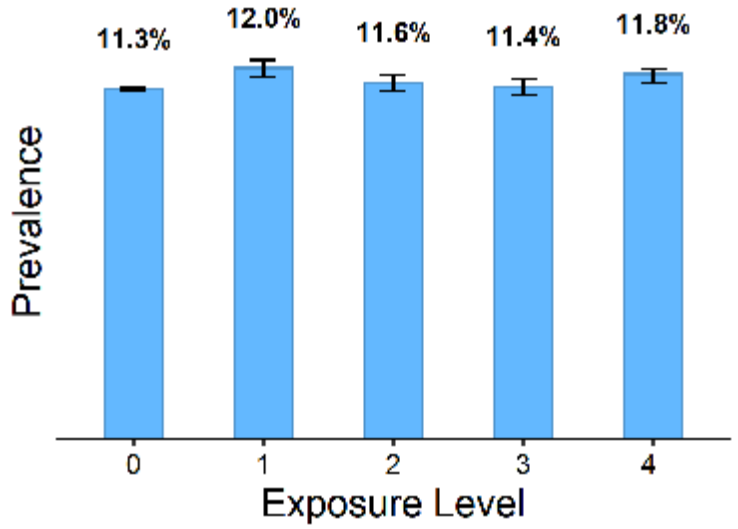

Figure 12. Crude prevalence of overweight by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

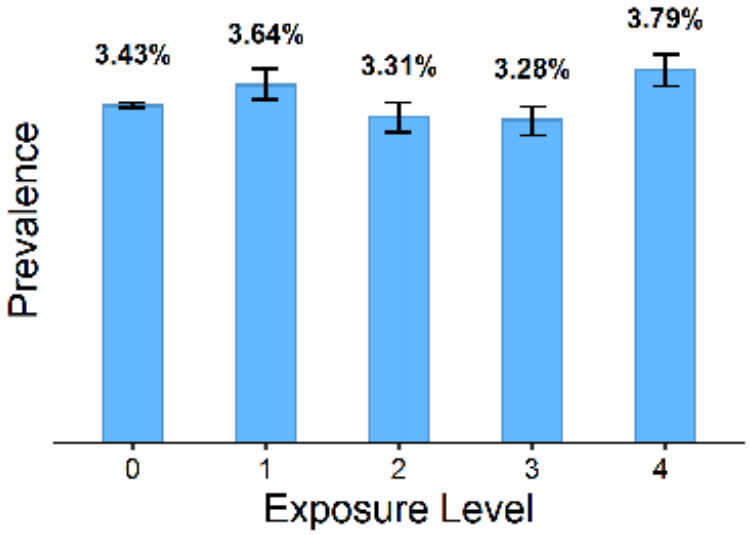

Figure 13. Crude prevalence of obesity by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

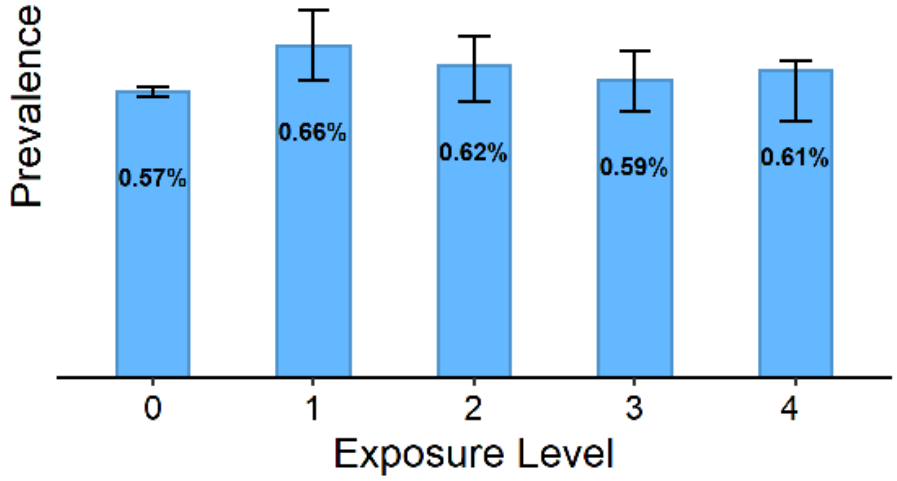

Figure 14. Crude prevalence of hypertension by HBA-IAP categories.
0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Adjusted associations of HBA-IAP with prevalent health conditions at age 17

Asthma

Multivariable logistic regression models examined adjusted associations between HBA-IAP and asthma. The basic model is adjusted for SES, year of birth, and school orientation (Figure 15). These results show that residency in HBA is associated with a higher risk of asthma (in comparison to non-HBA residency), but this association is limited to the three lowest HBA-IAP exposure categories. When examining the shape of the exposure-response curve, one can see that among categories 2-4, higher exposure is actually associated with lower adjusted risk.

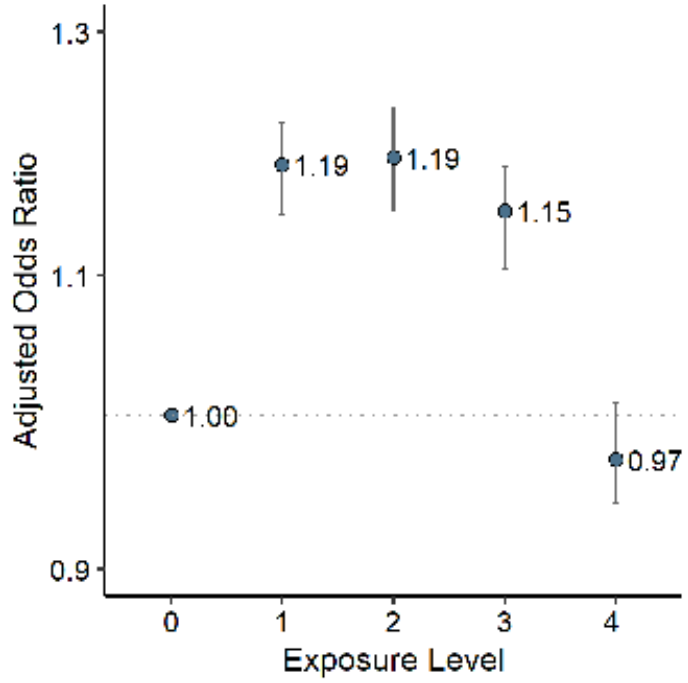

Figure 15. Association between HBA-IAP and asthma, adjusted for SES, year of birth and school orientation. N = 2,402,894, N cases = 144,358. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

We have also compared this curve to a similar curve that we get for our study population regarding known hazardous air pollutants. Figures 16 and 17 present adjusted associations of asthma with $PM_{2.5}$ and for NOx, respectively, and unlike Figure 15 – they demonstrate a clear classical exposure-response curve.

Figure 18 presents associations of HBA-IAP with asthma that are further adjusted for NOx, and Figure 19 adds adjustment for $PM_{2.5}$. These adjustments do not change the exposure-response curve substantially. The adjustment for $PM_{2.5}$ shows that the highest HBA-IAP exposure category is also at higher risk compared to non-HBA residents, but the shape of the curve within the three highest HBA-IAP exposure groups is still decreasing, such that higher exposure is associated with lower risk. One should also note that, as explained above, all models that are adjusted for $PM_{2.5}$ use a subset of ~35% of the study population because of the lack of available PM data for early decades.

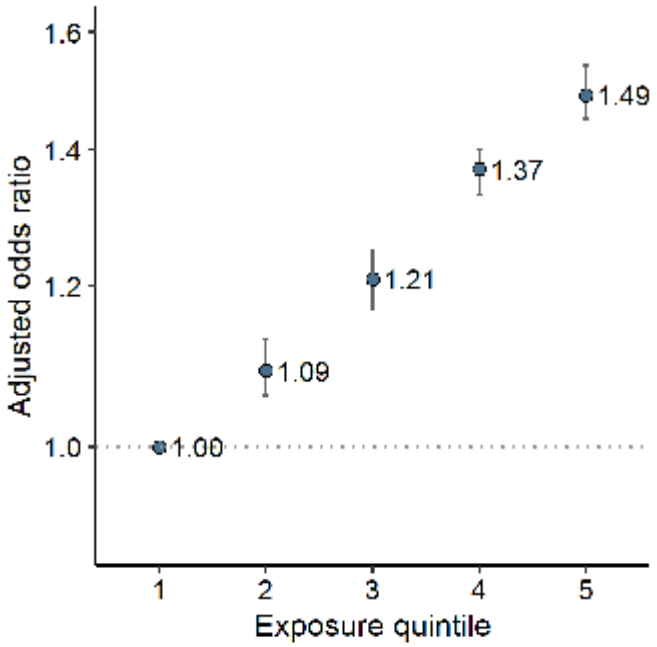

Figure 16. Association between $PM_{2.5}$ and asthma, adjusted for SES, year of birth and school orientation. N = 795,863, N cases = 57814. 1 = Reference category (lowest $PM_{2.5}$ quintile).

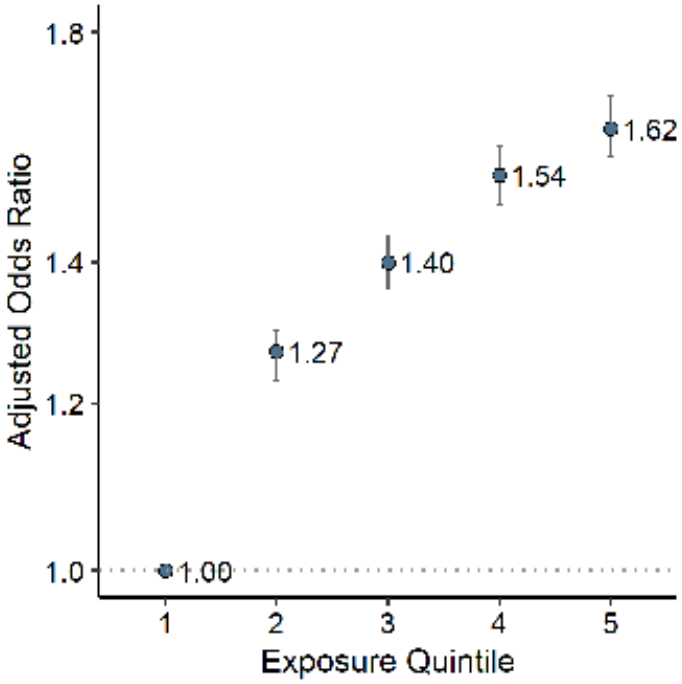

Figure 17. Association between NOx and asthma, adjusted for SES, year of birth and school orientation. N = 2,311,240, N cases = 140,166. 1 = Reference category (lowest NOx quintile).

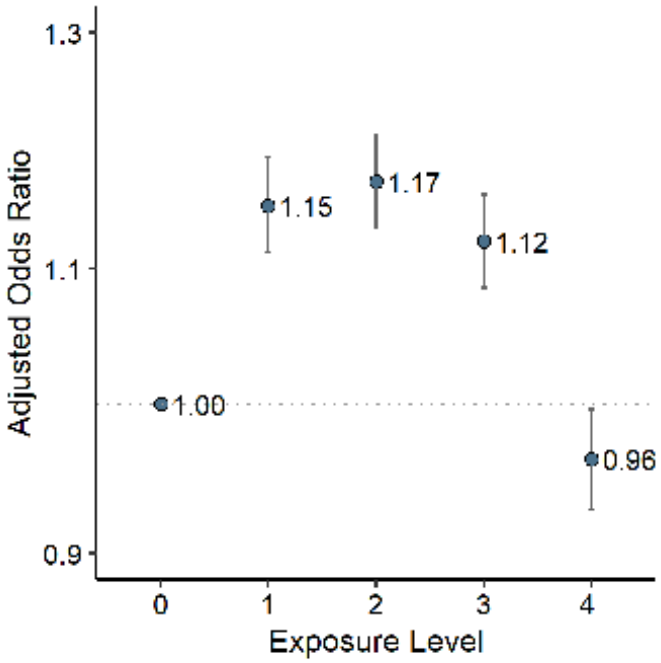

Figure 18. Association between HBA-IAP and asthma, adjusted for SES, year of birth, school orientation and NOx. N = 2,311,240, N cases = 140,166. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

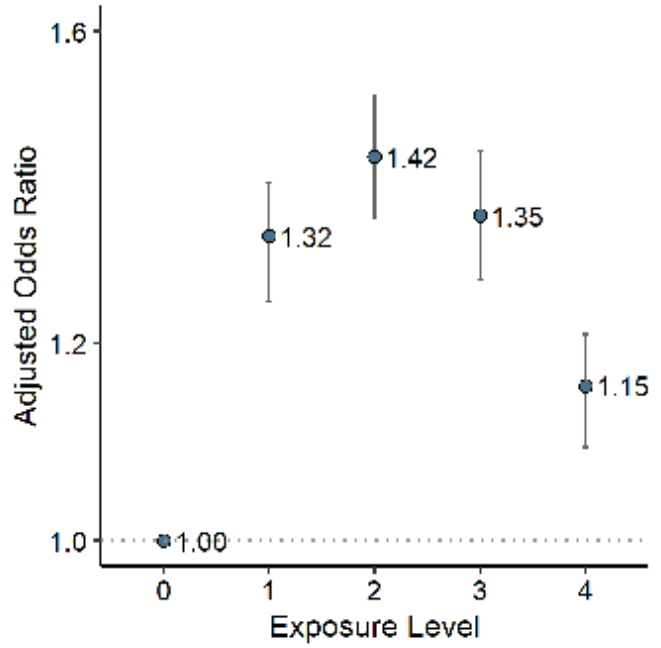

Figure 19. Association between HBA-IAP and asthma, adjusted for SES, year of birth, school orientation, NOx and $PM_{2.5}$. N = 791,132, N cases = 57,573. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Stratification by decades provides a different method to handle potential confounding by time trends. It is also important since the exposure model lacks any temporal variability (i.e., HBA-IAP exposures assessed by the exposure model are constant with time per any specific address). Since HBA-IAP has changed substantially over the 5- years exposure periods that are relevant to our study population, it is possible that this exposure was associated with higher risk of prevalent asthma at age 17 only in some of the historical decades we examined, but not in others. Figure 20 shows associations between HBA-IAP exposure and asthma, stratified by four periods, based on years of birth of the study population. Despite some differences in the magnitude of the associations among periods, and some differences in the shape of the exposure-response curves, the highest exposure category is associated with the lowest adjusted risk of asthma among all HBA groups in all periods. Further adjustment for NOx and $PM_{2.5}$ does not change this general picture, as can be seen in Figures 21 and 22, respectively.

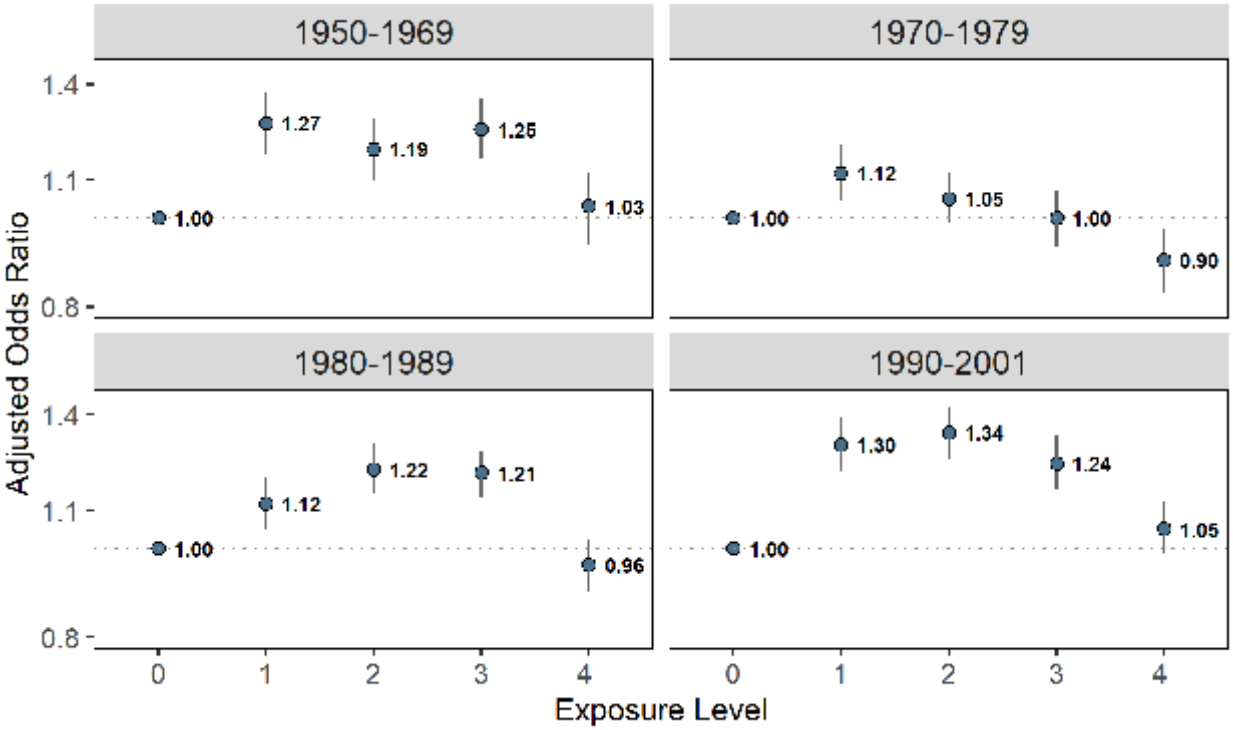

Figure 20. Association between HBA-IAP and asthma, stratified by decade of birth and adjusted for SES, year of birth and school orientation. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

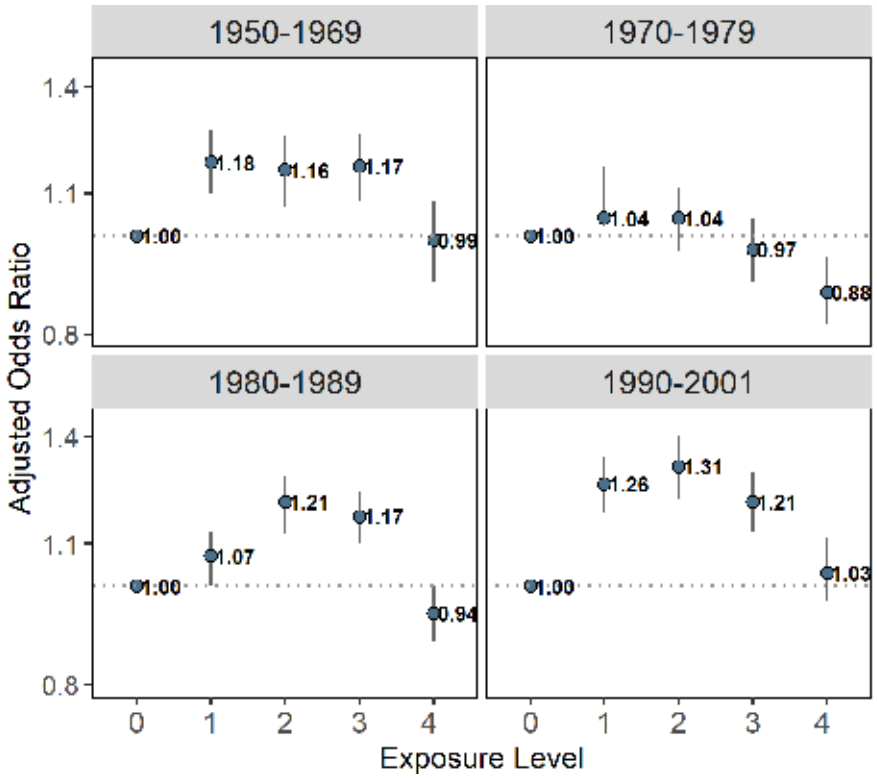

Figure 21. Association between HBA-IAP and asthma, stratified by decade of birth and adjusted for SES, year of birth, school orientation and NOx. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

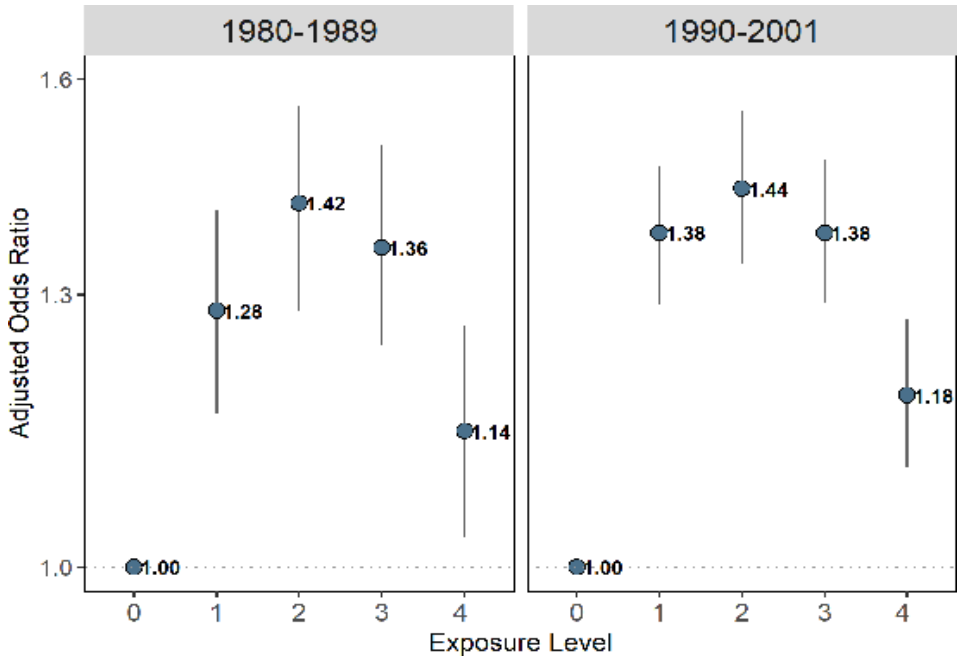

Figure 22. Association between HBA-IAP and asthma, stratified by decade of birth and adjusted for SES, year of birth, school orientation, NOx and $PM_{2.5}$. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Further stratification by parental country of origin was performed to partly exclude potential cultural and genetic confounders. We have defined a short list of countries that are commonly known as Jewish "Ashkenazi" origin (Russia and other former USSR countries, Poland, Hungary, Romania, Germany) and another list of countries that are commonly known as Jewish "Mizrahi" origin (Morocco, Iraq, Egypt, Turkey, Yemen and Iran). Figure 23 shows adjusted associations stratified by six groups created by parental origin (see figure caption for more details). As can be seen from this figure, despite some differences, especially between the two lowest HBA-IAP categories within the HBA population, all strata show quite similar exposure-response curves which are also compliant with the previous, unstratified models.

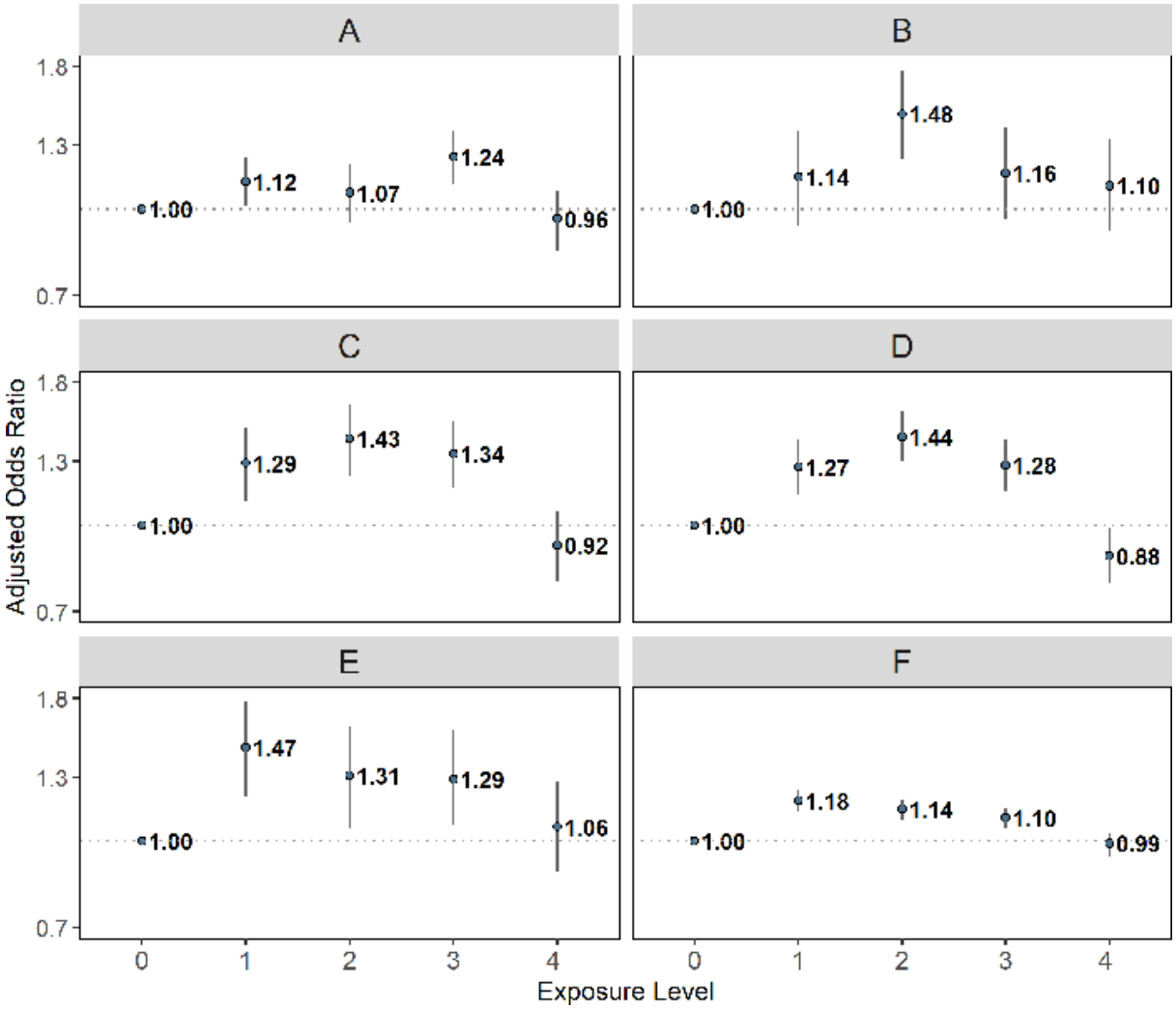

Figure 23. Association between HBA-IAP and asthma, stratified by parental country of birth and adjusted for SES, year of birth, and school orientation. A = both parents from Jewish "Ashkenazi" country (N = 147,121); B = one parent from Jewish "Ashkenazi" country and the other one Israeli-born (N = 61,636); C = one parent from Jewish "Ashkenazi" country and the other one from Jewish "Mizrahi" country (N = 86,759); D = both parents from Jewish "Mizrahi" country (N = 321,001); E = one parent from Jewish " Mizrahi" country and the other one Israeli-born (N = 79,002); F = all other participants (N = 1,828,226). 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

We have also considered two additional categorizations of HBA-IAP by equal exposure intervals instead of equal participant numbers. Table 5 and Figure 24 represent the categorization of HBA-IAP into four equal interval exposure groups. Figures 25 and 26 display the prevalence of asthma and the adjusted associations by these categories, respectively.

Table 5. Equal-interval HBA-IAP exposure (four groups).

HBA-IAP category	N
Reference (non-HBA)	2,216,927
1	63,298
2	140,262
3	20,301
4	2,437

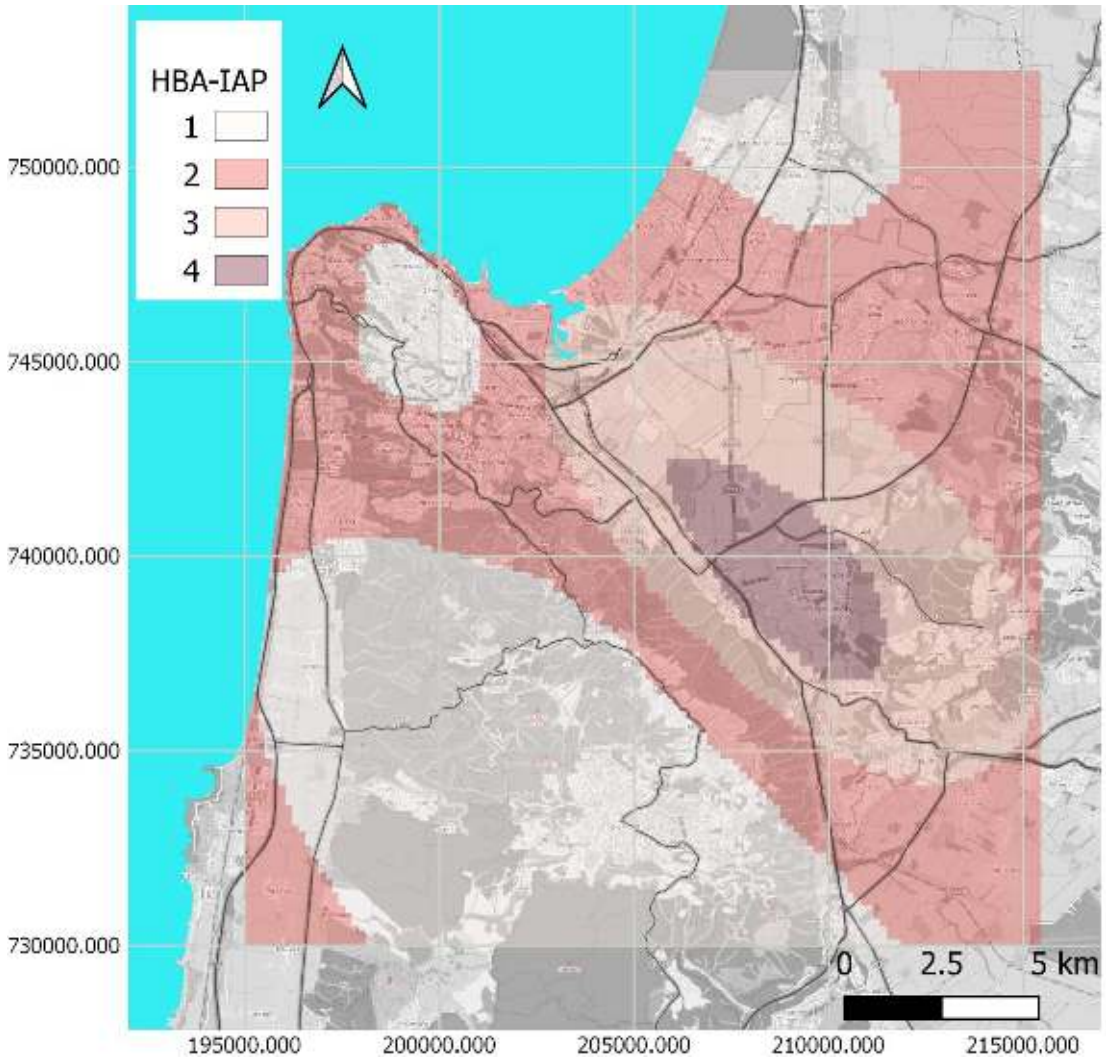

Figure 24. A map of the HBA-IAP exposure model, with equal-interval color-coded exposure categories 1-4.

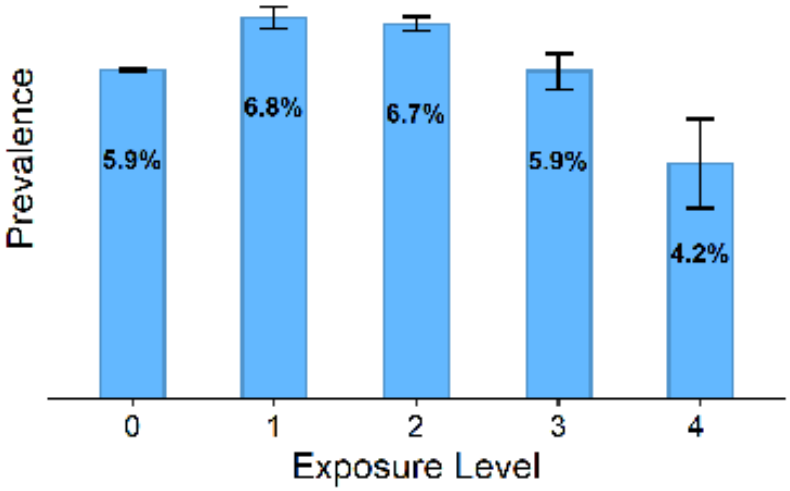

Figure 25. Crude prevalence of Asthma by HBA-IAP four equal-interval categories. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

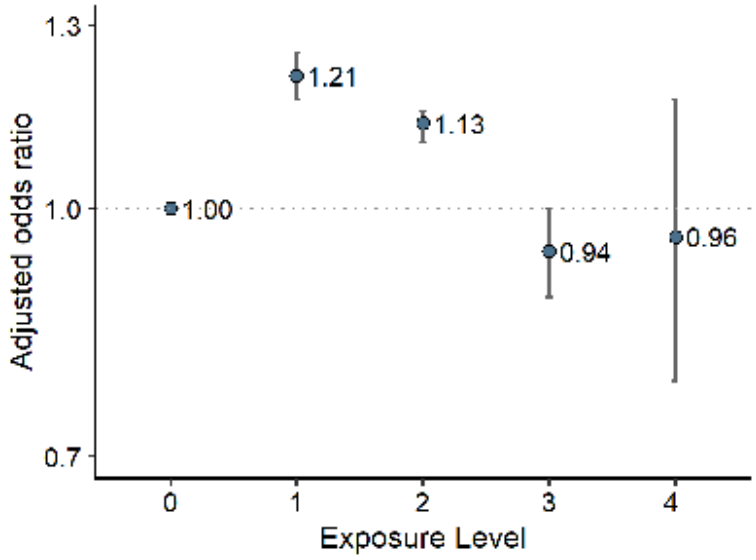

Figure 26. Association between HBA-IAP and asthma with four equal-interval categories, adjusted for SES, year of birth and school orientation. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Table 6 and Figure 27 represent HBA-IAP categorization into three equal interval exposure groups. Figures 28 and 29 display the prevalence of asthma and the adjusted associations by these categories, respectively. These two additional categorizations of the exposure give a similar picture to our main categorization in terms of the exposure-response curve.

Table 6. Equal-interval HBA-IAP exposure (three groups).

HBA-IAP category	N
Reference (non-HBA)	2,216,927
1	130,156
2	92,520
3	3,640

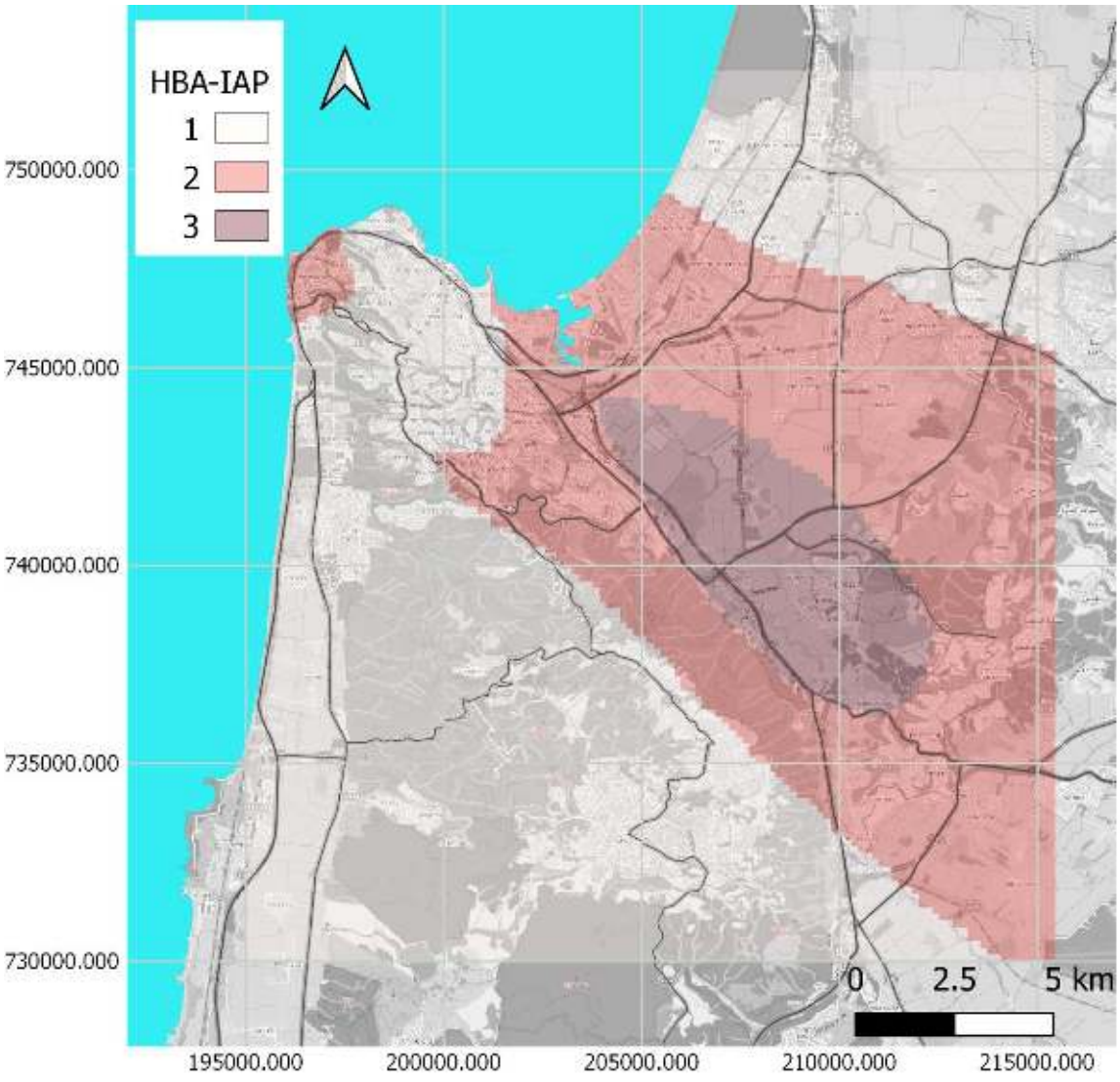

Figure 27. A map of the HBA-IAP exposure model, with equal-interval color-coded exposure categories 1-3.

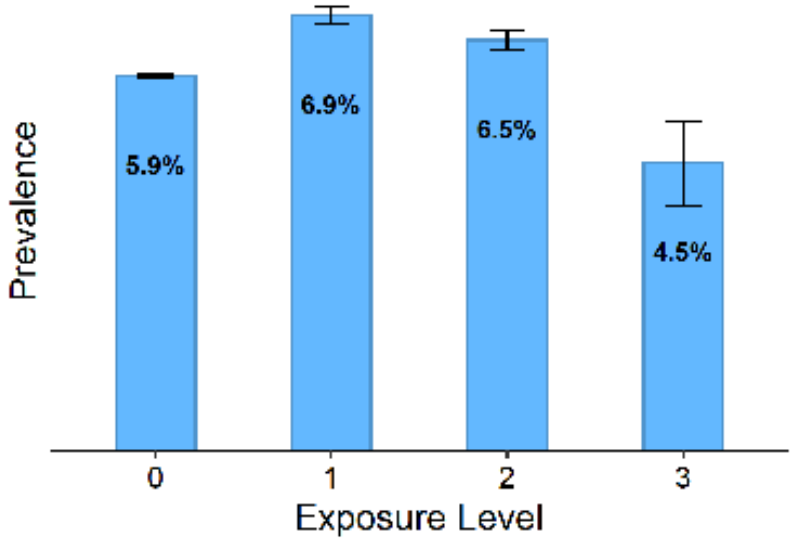

Figure 28. Crude prevalence of Asthma by HBA-IAP three equal-interval categories. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

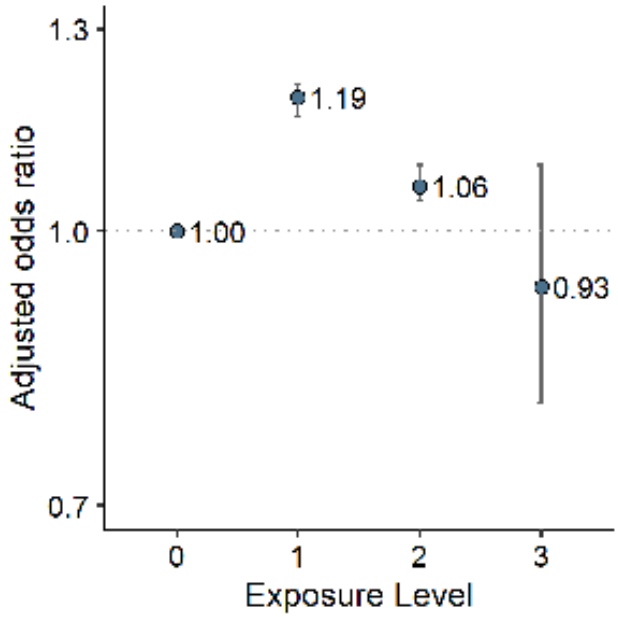

Figure 29. Association between HBA-IAP and asthma with three equal-interval categories, adjusted for SES, year of birth and school orientation. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Other atopic diseases

Examination of associations between HBA-IAP and other prevalent atopic diseases at age 17: rhinitis (with or without comorbid asthma), rhinoconjunctivitis, and atopic dermatitis, showed similar exposure-response curves, where residency in HBA is associated with higher risk, except for the highest exposure group (Figure 30). Further adjustment for NOx and $PM_{2.5}$ did not change this picture materially (Figures 31 and 32, respectively).

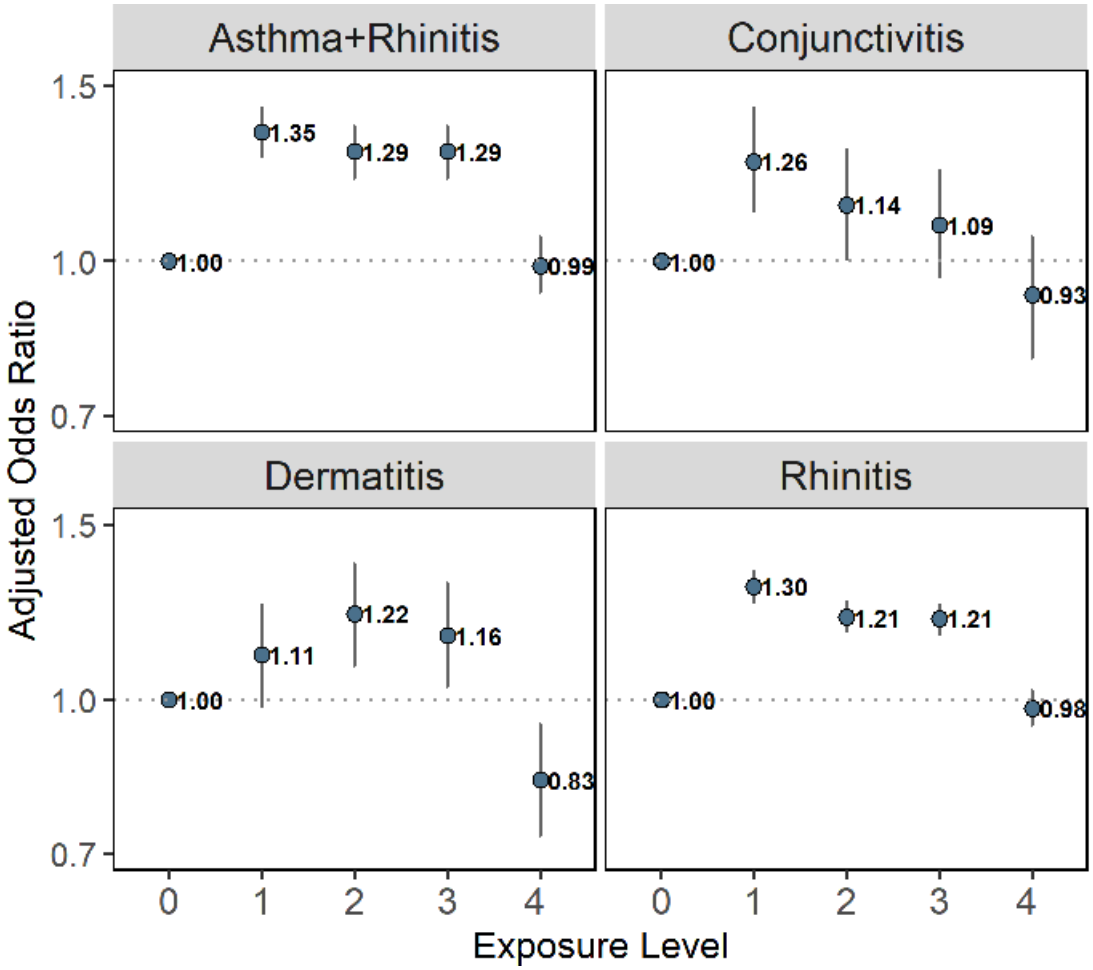

Figure 30. Association between HBA-IAP and other atopic diseases, adjusted for SES, year of birth, and school orientation. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

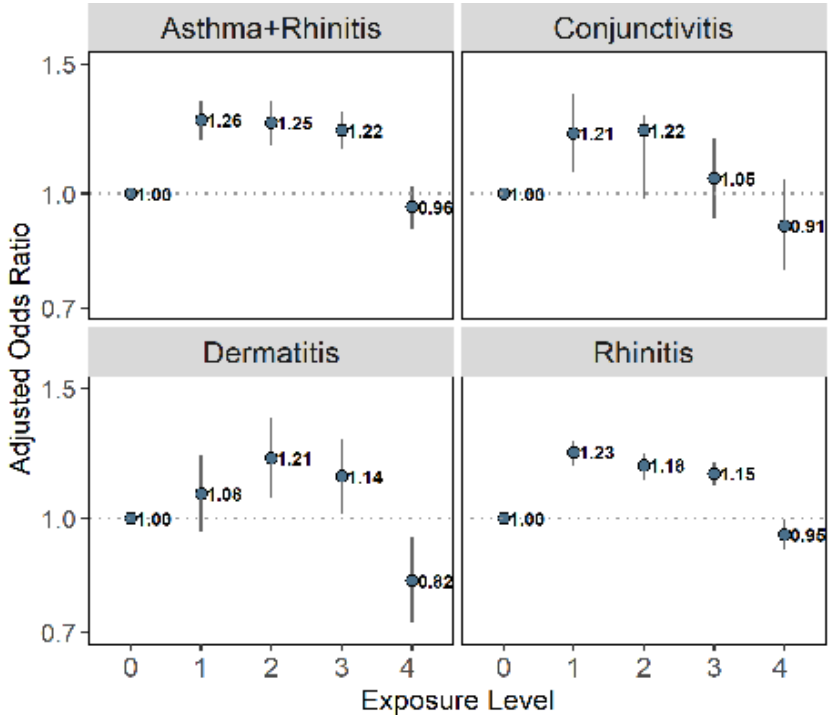

Figure 31. Association between HBA-IAP and other atopic diseases, adjusted for SES, year of birth, school orientation, and NOx. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

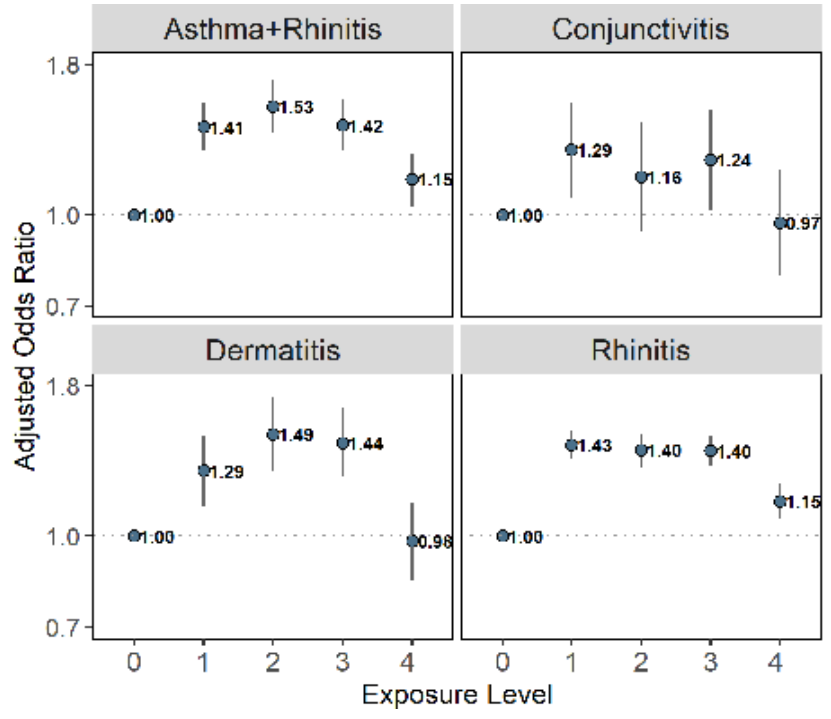

Figure 32. Association between HBA-IAP and other atopic diseases, adjusted for SES, year of birth, school orientation, NOx and $PM_{2.5}$. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Metabolic outcomes

Figures 33 and 34 examine HBA-IAP's adjusted associations with BMI and hypertension, respectively. Obesity and overweight were associated with HBA residency, without a clear exposure-response curve, while the HBA population had a lower adjusted risk of being underweight. Hypertension was not associated with HBA-IAP.

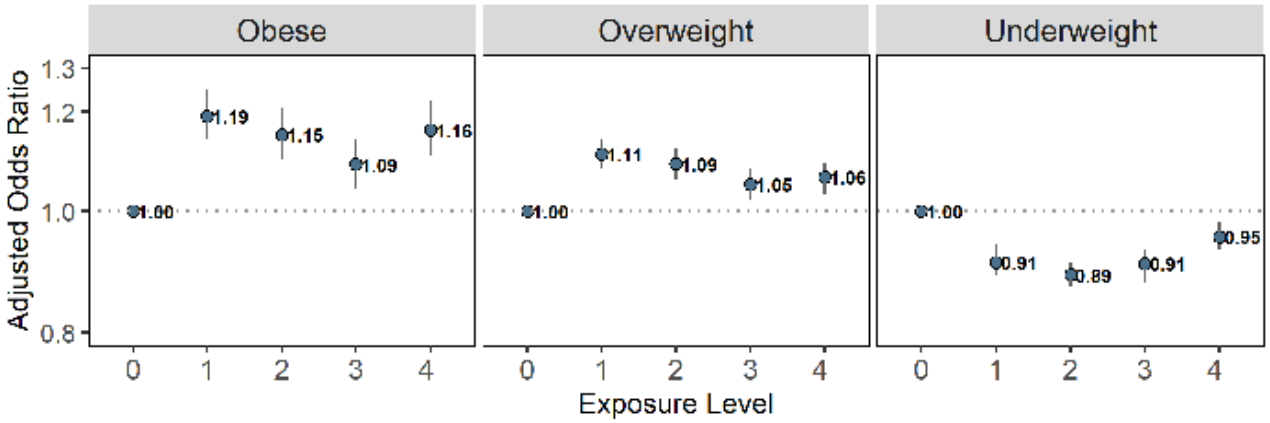

Figure 33. Association between HBA-IAP and BMI categories (taken from a multinomial regression model), adjusted for SES, year of birth, and school orientation. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

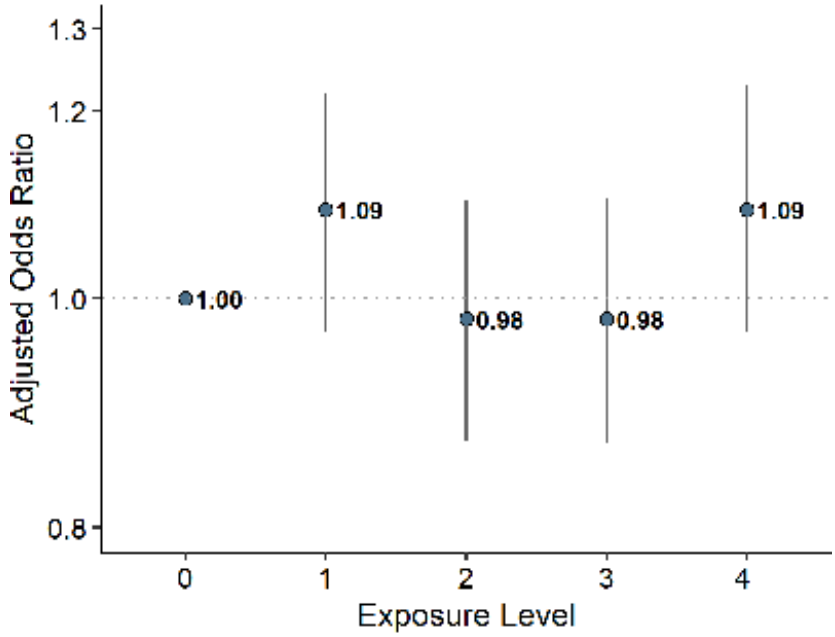

Figure 34. Association between HBA-IAP and hypertension, defined as SBP>140 and DBP>90, adjusted for SES, year of birth, and school orientation. 0 = Reference category (non-HBA residents), 1 = lowest HBA-IAP exposure.

Cancer

Study population

The cohort included N=2,187,317 participants, of which 59% were males. The participants contributed 41,696,278 person-years of follow-up. Table S7 and Figure S1 represent HBA-IAP categorization into the three equal-participants exposure groups that were used in the cancer analyses. Over 45 years, 47,129 participants were diagnosed with cancer, with a crude incidence rate of 142.8 cases per 100,000 person-years in the non-exposed category and 171.5, 171.7, and 174.8 cases per 100,000 person-years in the low, intermediate and high HBA-IAP exposure levels, respectively (Tables 7a, 7b).

The most common cancer in the population was female breast cancer, with a crude incidence rate of 26.0 cases per 100,000 person-years in the non-exposed category and 31.0, 31.9, and 31.0 cases per 100,000 person-years in the low, intermediate and high exposure groups, respectively. The distribution of the general population and cancer diagnoses by exposure levels and subject characteristics are presented in Tables 7a, 7b.

Table 7a. Descriptive statistics of the study population by cancer category (N = 2,187,317).

Characteristic	Entire Cohort	Any Cancer	Head & Neck	GI	Pulmonary	Skin	Breast (Female)	Female Reproductive
Total N	2,187,317	47,129	1292	4103	1925	5185	8576	2038
Sex								
Male	1,294,570 (59.2)	21,097 (44.8)	1005 (77.8)	2764 (67.4)	1444 (75.0)	2785 (53.7)	-	-
Female	892,747 (40.8)	26,032 (55.2)	287 (22.2)	1339 (32.6)	481 (25.0)	2400 (46.3)	8,576	2,038
Year of birth								
1947-1959	320,470 (14.7)	22,154 (47.0)	782 (60.5)	2699 (65.8)	1374 (71.4)	2549 (49.2)	4410 (51.4)	965 (47.4)
1960--1970	440,294 (20.1)	13,990 (29.7)	329 (25.5)	985 (24.0)	430 (22.3)	1565 (30.2)	3088 (36.0)	669 (32.8)
1971-1980	571,908 (26.1)	8269 (17.5)	120 (9.3)	363 (8.8)	100 (5.2)	860 (16.6)	985 (11.5)	322 (15.8)
1981-2001	854,646 (39.1)	2716 (5.8)	61 (4.7)	56 (1.4)	21 (1.1)	211 (4.1)	93 (1.1)	82 (4.0)
SES Z score								
1st	425,733 (19.5)	7810 (16.6)	315 (24.4)	836 (20.4)	459 (23.8)	531 (10.2)	1010 (11.8)	303 (14.9)
2nd	429,549 (19.6)	9166 (19.4)	276 (21.4)	850 (20.7)	407 (21.1)	844 (16.3)	1562 (18.2)	389 (19.1)
3rd	422,347 (19.3)	8999 (19.1)	232 (18.0)	771 (18.8)	363 (18.9)	928 (17.9)	1672 (19.5)	397 (19.5)
4th	416,107 (19.0)	9933 (21.1)	237 (18.3)	799 (19.5)	374 (19.4)	1226 (23.6)	1957 (22.8)	469 (23.0)
5th	417,808 (19.1)	10,479 (22.2)	218 (16.9)	769 (18.71)	295 (15.3)	1562 (30.1)	2264 (26.4)	454 (22.3)
Missing	75,774 (3.5)	742 (1.6)	14 (1.1)	78 (1.9)	27 (1.4)	94 (1.8)	111 (1.3)	26 (1.3)
IAP-HBA exposure, N(%)								
Reference	1,945,991 (90.4)	41,420 (87.9)	1145 (88.6)	3673 (89.5)	1713 (89.0)	4429 (85.4)	7540 (87.9)	1792 (87.9)
1	68,587 (3.2)	1955 (4.1)	51 (3.9)	156 (3.8)	81 (4.2)	249 (4.8)	353 (4.1)	94 (4.6)
2	69,149 (3.2)	1957 (4.2)	46 (3.6)	136 (3.3)	70 (3.6)	270 (5.2)	364 (4.2)	80 (3.9)
3	67,871 (3.2)	1797 (3.8)	50 (3.9)	138 (3.4)	61 (3.2)	237 (4.6)	319 (3.7)	72 (3.5)
IAP-HBA exposure, Incidence rate (per 100,000 PY)								
0		142.8	3.9	12.7	5.9	15.3	26.0	6.2
1		171.5	4.5	13.7	7.1	21.9	31.0	8.3
2		171.7	4.0	11.9	6.2	23.7	31.9	7.0
3		174.8	4.9	13.4	5.9	23.1	31.0	7.0
Cognitive score								
Low	641,826 (30.2)	9161 (19.6)	348 (27.4)	948 (23.5)	603 (32.0)	517 (9.9)	872 (10.2)	349 (17.2)

Intermediate	719,878 (33.9)	11,296 (24.2)	312 (24.6)	865 (21.5)	390 (20.7)	1074 (20.8)	1686 (19.7)	484 (23.8)
High	462,524 (21.8)	16,919 (36.3)	430 (33.9)	1566 (38.9)	654 (34.8)	2018 (39.1)	3830 (44.8)	846 (41.7)
Very high	300,299 (14.1)	9266 (19.9)	179 (14.1)	651 (16.2)	235 (12.5)	1552 (30.1)	2156 (25.2)	350 (17.2)

Table 7b. Descriptive statistics of the study population by cancer category (N = 2,187,317).

Characteristic	Entire Cohort	Male reproductive	Urinary	CNS	Thyroid	Hodgkin's Lymphoma	NHL	Leukemias
Total N	2,187,317	2974	2325	1357	2899	1872	3301	1597
Sex								
Male	1,294,570 (59.2)	2974	1979 (85.1)	1268 (58.8)	838 (28.9)	1066 (56.9)	2234 (67.7)	1107 (69.3)
Female	892,747 (40.8)		346 (14.9)	890 (41.2)	2,061 (71.1)	806 (43.1)	1,067 (32.2)	490 (30.7)
Year of birth								
1947-1959	320,470 (14.7)	1613 (54.2)	1529 (65.8)	922 (42.7)	853 (29.4)	382 (20.4)	1543 (46.7)	779 (48.8)
1960--1970	440,294 (20.1)	569 (19.1)	576 (24.8)	618 (28.6)	987 (34.0)	537 (28.7)	1010 (30.6)	430 (26.9)
1971-1980	571,908 (26.1)	552 (18.6)	181 (7.8)	423 (19.6)	799 (27.6)	642 (34.3)	551 (16.7)	273 (17.1)
1981-2001	854,646 (39.1)	240 (8.1)	39 (1.7)	195 (9.0)	260 (9.0)	311 (16.6)	197 (6.0)	115 (7.2)
SES Z score								
1st	425,733 (19.5)	507 (17.0)	460 (19.8)	429 (19.9)	438 (15.1)	361 (19.3)	665 (20.1)	342 (21.4)
2nd	429,549 (19.6)	568 (19.1)	468 (20.1)	420 (19.5)	599 (20.7)	366 (19.5)	675 (20.4)	319 (20.0)
3rd	422,347 (19.3)	615 (20.7)	472 (20.3)	410 (19.0)	591 (20.4)	351 (18.8)	603 (18.3)	302 (18.9)
4th	416,107 (19.0)	609 (20.5)	466 (20.0)	441 (20.4)	592 (20.4)	368 (19.7)	670 (20.3)	314 (19.7)
5th	417,808 (19.1)	628 (21.1)	416 (17.9)	406 (18.8)	637 (22.0)	399 (21.3)	639 (19.4)	288 (18.0)
Missing	75,774 (3.5)	47 (1.6)	43 (1.8)	52 (2.4)	42 (1.4)	27 (1.4)	49 (1.5)	32 (2.0)
IAP-HBA exposure, N(%)								
Reference	1,945,991 (90.4)	2631 (88.5)	2103 (90.5)	1887 (87.4)	2543 (87.7)	1663 (88.8)	2942 (89.1)	1416 (88.7)
1	68,587 (3.2)	108 (3.6)	86 (3.7)	89 (4.1)	98 (3.4)	76 (4.1)	140 (4.2)	59 (3.7)
2	69,149 (3.2)	129 (4.3)	71 (3.1)	93 (4.3)	133 (4.6)	65 (3.5)	100 (3.0)	54 (3.4)
3	67,871 (3.2)	106 (3.6)	65 (2.8)	89 (4.1)	125 (4.3)	68 (3.6)	119 (3.6)	68 (4.3)
IAP-HBA exposure, Incidence rate (per 100,000 PY)								
0		9.1	7.2	6.5	8.8	5.7	10.1	4.9
1		9.5	7.6	7.8	8.6	6.7	12.3	5.2
2		11.3	6.2	8.2	11.7	5.7	8.8	4.7
3		10.3	6.3	8.7	12.2	6.6	11.8	6.6
Cognitive score								
Low	641,826 (30.2)	594 (20.2)	581 (25.3)	524 (24.8)	538 (18.7)	496 (26.8)	727 (22.3)	394 (25.4)

Intermediate	719,878 (33.9)	742 (25.3)	510 (22.2)	571 (27.0)	831 (28.8)	577 (31.2)	784 (24.0)	384 (24.7)
High	462,524 (21.8)	960 (32.7)	835 (36.4)	635 (30.0)	944 (32.8)	485 (26.2)	1152 (35.3)	529 (34.0)
Very high	300,299 (14.1)	638 (21.7)	368 (16.0)	387 (18.3)	569 (19.7)	294 (15.9)	597 (18.3)	246 (15.8)

Associations between cancer and HBA-IAP exposure

Any cancer

In unadjusted Cox models, we observed clear positive exposure-response associations between HBA-IAP exposure and any cancer, with an HR of 1.23, 95% CI = 1.17 to 1.29 for the highest exposure category, compared to the reference category (non-HBA residents). The association slightly weakened when adjusting for potential confounders, with an HR of 1.16, 95% CI = 1.10 to 1.21 for the highest exposure category (Figure 35, Table 8).

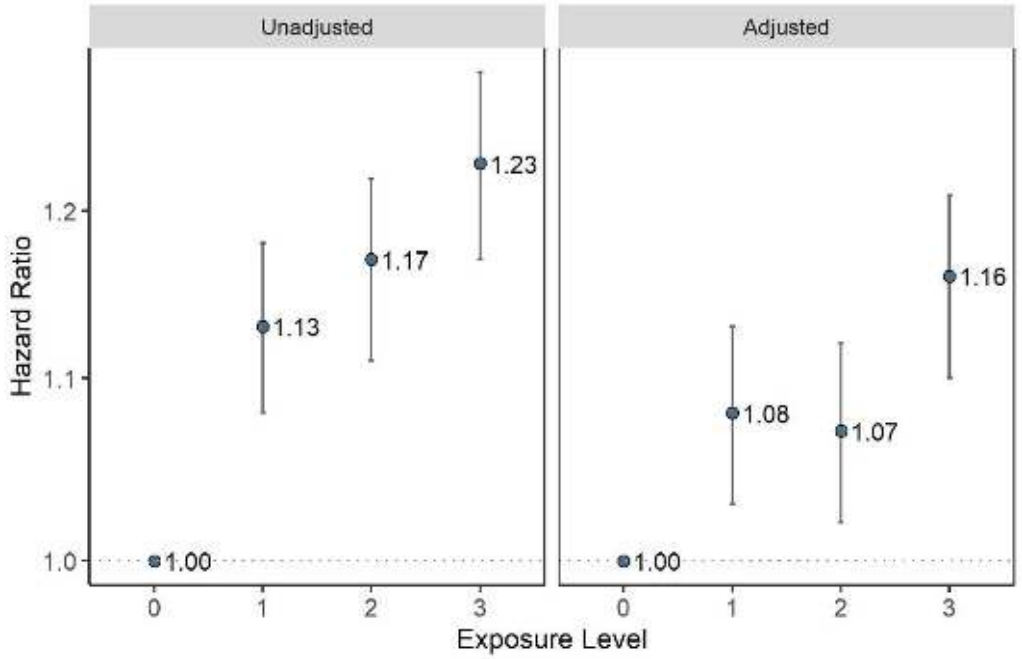

Figure 35. Associations of exposure to HBA-IAP with any-cancer, modeled using Cox proportional hazard regression. 0 = reference category (non-HBA residents), 1 = low HBA-IAP exposure level, 2= intermediate exposure, 3= high exposure. Left panel: unadjusted associations. N= 2,151,598, events= 47,129. Right panel: associations adjusted for sex, year of birth, locality type, origin, cognitive score and NOx. N=2,022,930, events= 45,141.

Table 8. Unadjusted and adjusted associations of exposure to HBA-IAP with any-cancer.

HBA-IAP	Low	Intermediate	High	No. of events
Unadjusted model	1.13	1.17	1.23	47,129
	(1.08 to1.18)	(1.11 to 1.22)	(1.17 to 1.29)	
Adjusted model	1.08	1.07	1.16	45,141
	(1.03 to 1.13)	(1.02 to 1.12)	(1.10 to 1.21)	

Analyses were adjusted by sex, year of birth, locality type, origin and cognitive score. *HR=Hazard Ratio, CI= confidence interval*

When stratifying the study population by decade of birth, we found monotonic exposure-response curves for the association between level of exposure and any cancer at all periods, excluding 1981-2001. The magnitude of the associations was stronger for those born after 1970 (Figure 36, Table 9).

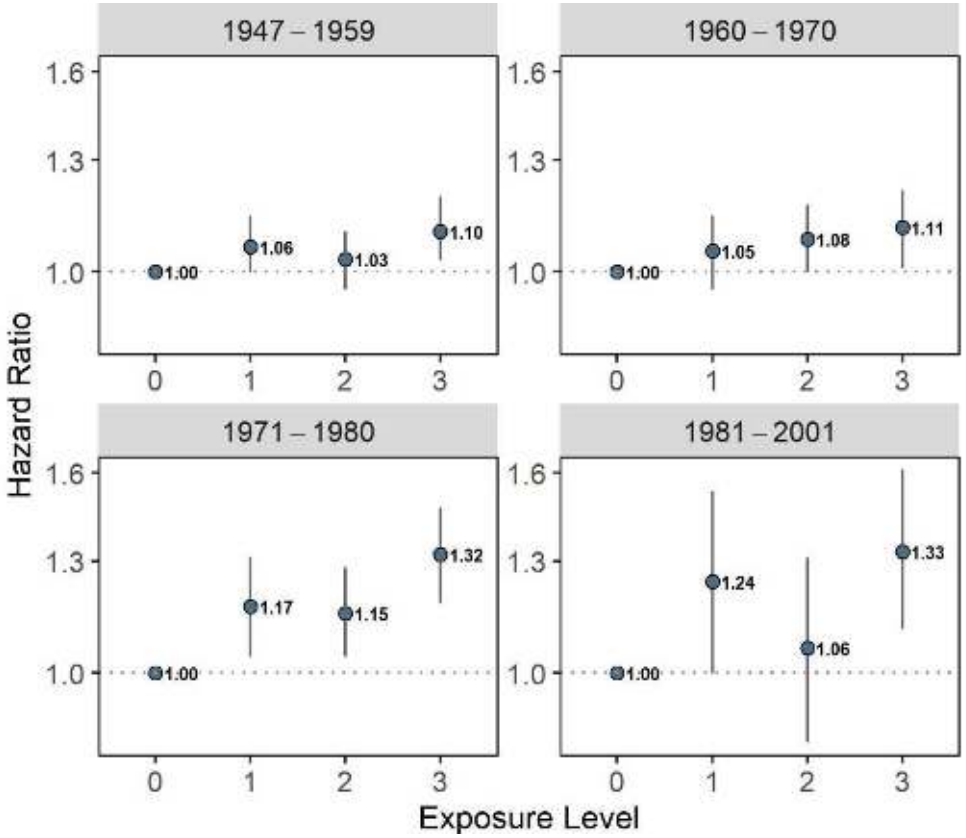

Figure 36. Associations of exposure to HBA-IAP with any-cancer, modeled using Cox proportional hazard regression, stratified by the decade of birth and adjusted for sex, locality type, origin and cognitive score. 1947-1959: *N*=278,301, cases=21,068; 1960-1970: *N*=421,408, cases=13,628; 1971-1980: *N*=543,818, cases=7954; 1981-2001: *N*=779,403, cases=2491. Reference category: non-HBA residents.

Table 9. Associations of exposure to HBA-IAP with any-cancer, modeled using Cox proportional hazard regression, stratified by the decade of birth and adjusted for sex, locality type, origin and cognitive score.

HBA-IAP	Low	Intermediate	High
Decade of Birth			
1947-1959	1.06	1.03	1.10
	(1.00 to 1.14)	(0.96 to 1.10)	(1.03 to 1.19)
1960-1970	1.05	1.08	1.11
	(0.964 to 1.14)	(1.00 to 1.17)	(1.01 to 1.21)
1971-1980	1.17	1.15	1.32
	(1.04 to 1.31)	(1.04 to 1.28)	(1.18 to 1.47)
1981-2001	1.24	1.06	1.33
	(1.00 to 1.53)	(0.85 to 1.31)	(1.11 to 1.61)

When stratifying the study population by SES quintiles, we found monotonic exposure-response curves in most SES quintiles, with a stronger association in the lower SES quintile than in all other quintiles (Figure 37, Table 10).

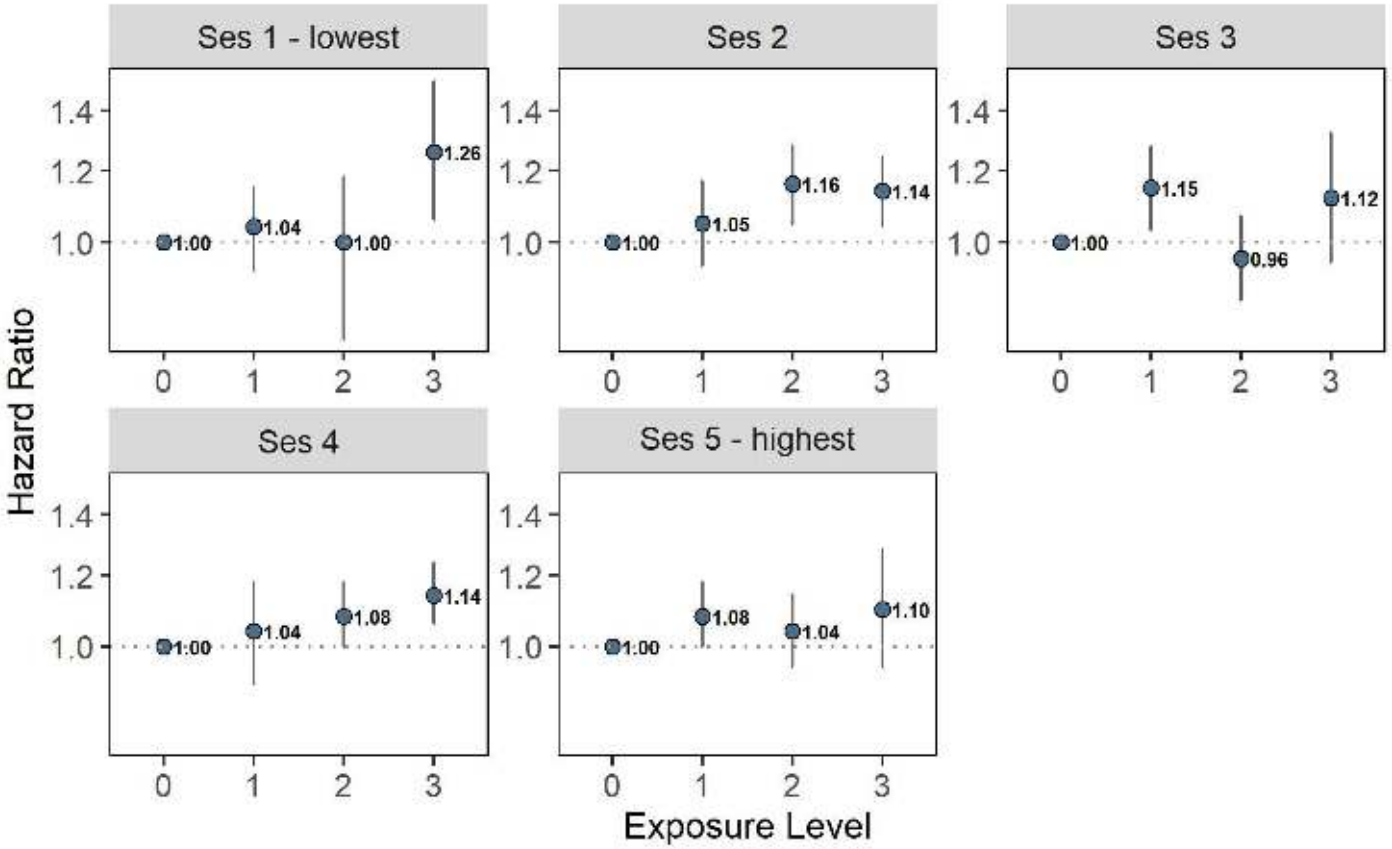

Figure 37. Effect modification by SES for the association between HBA-IAP and any-cancer, modeled using Cox proportional hazard regression, adjusted for sex, year of birth, locality type, origin, cognitive score and NOx. Higher scores in the census SES index represent higher socioeconomic status. SES 1: *N*=425,733, cases=7522; SES 2: *N*=429,549, cases=8505; SES 3: *N*=422,347, cases=8711; SES 4: *N*=416,107, cases=9631; SES 5: *N*=417,808, cases=10,320. 0 = reference category (non-HBA residents), 1 = low HBA-IAP exposure level, 2= intermediate exposure, 3= high exposure.

Table 10. Effect modification by SES for the association between HBA-IAP and any-cancer, modeled using Cox proportional hazard regression, adjusted for sex, year of birth, locality type, origin, cognitive score and NOx.

HBA-IAP	Low	Intermediate	High
SES			
SES 1 - lowest	1.04	1.0	1.26
	(0.93 to 1.15)	(0.78 to 1.18)	(1.06 to 1.51)
SES 2	1.05	1.16	1.14
	(0.94 to 1.17)	(1.05 to 1.28)	(1.04 to 1.24)
SES 3	1.15	0.96	1.12
	(1.03 to 1.28)	(0.86 to 1.07)	(0.95 to 1.32)
SES 4	1.04	1.08	1.14
	(0.91 to 1.18)	(1.00 to 1.18)	(1.06 to 1.24)
SES 5 - highest	1.08	1.04	1.10
	(1.00 to 1.18)	(0.95 to 1.14)	(0.95 to 1.28)

Specific cancer categories

When examining specific cancers, unadjusted Cox models show monotonic exposure-response associations for several cancer groups. The unadjusted HRs in the highest HBA-IAP exposure category compared to the reference category was 1.52 (95% CI = 1.33 to 1.73) for melanoma; 1.36 (95% CI = 1.07 to 1.73) for leukemia; 1.39 (95% CI = 1.16 to 1.66) for thyroid cancer; 1.33 (95% CI = 1.07 to 1.64) for CNS tumors; 1.20 (95% CI = 1.08 to 1.35) for head and neck cancers, and 1.20 (95% CI = 1.08 to 1.35) for female breast cancer (Table 11).

The adjusted associations between exposure to HBA-IAP and specific cancers are displayed in Table 11 and Figure 37. The associations demonstrated positive monotonic or nearly-monotonic exposure-response curves for 6 out of 13 cancer groups: female breast cancer, CNS, head and neck, leukemia, melanoma, and thyroid cancer. On the other hand, the following cancer categories were not consistently associated with HBA-IAP exposure: gastrointestinal organs, Hodgkin's lymphoma, NHL, reproduction organs (in males and females), urinary tract, and pulmonary cancer.

Table 11. Associations between HBA-IAP exposure and specific cancers.

	Unadjusted			Adjusted		
HBA-IAP	**Low**	**Intermediate**	**High**	**Low**	**Intermediate**	**High**
Cancer type						
Head & Neck	1.06 (0.80 to 1.40)	0.99 (0.74 to 1.33)	1.24 (0.94 to 1.64)	1.05 (0.80 to 1.40)	0.95 (0.70 to 1.30)	1.22 (0.91 to 1.63)
GI	0.99 (0.84 to 1.16)	0.90 (0.76 to 1.10)	1.07 (0.90 to 1.27)	0.96 (0.81 to 1.12)	0.91 (0.76 to 1.08)	1.03 (0.86 to 1.23)
Pulmonary	1.09 (0.87 to 1.36)	0.99 (0.78 to 1.26)	1.00 (0.78 to 1.30)	1.10 (0.87 to 1.38)	1.06 (0.82 to 1.35)	1.01 (0.77 to 1.31)
Melanoma	1.35 (1.19 to 1.53)	1.51 (1.33 to 1.70)	1.52 (1.33 to 1.73)	1.15 (1.01 to 1.30)	1.17 (1.04 to 1.33)	1.28 (1.12 to 1.47)
Breast (Female)	1.10 (0.98 to 1.22)	1.18 (1.06 to 1.31)	1.20 (1.08 to 1.35)	1.02 (0.9 to 1.13)	1.04 (0.93 to 1.16)	1.14 (1.01 to 1.27)
Reproduction (Female)	1.25 (1.01 to 1.54)	1.10 (0.88 to 1.37)	1.14 (0.90 to 1.44)	1.19 (0.96 to 1.46)	1.03 (0.82 to 1.29)	1.12 (0.88 to 1.42)
Reproduction (Male)	0.98 (0.81 to 1.19)	1.20 (1.01 to 1.43)	1.14 (0.94 to 1.39)	0.94 (0.78 to 1.15)	1.12 (0.93 to 1.33)	1.12 (0.92 to 1.37)
Urinary	0.95 (0.76 to 1.18)	0.82 (0.65 to 1.04)	0.88 (0.69 to 1.13)	0.96 (0.77 to 1.20)	0.85 (0.67 to 1.08)	0.90 (0.70 to 1.16)
CNS	1.15 (0.93 to 1.42)	1.23 (1.00 to 1.51)	1.33 (1.07 to 1.64)	1.16 (0.93 to 1.44)	1.19 (0.96 to 1.48)	1.32 (1.06 to 1.64)
Thyroid	0.94 (0.77 to 1.15)	1.30 (1.09 to 1.55)	1.39 (1.16 to 1.66)	0.91 (0.74 to 1.11)	1.17 (0.98 to 1.40)	1.28 (1.07 to 1.54)
Hodgkin Lymphoma	1.17 (0.93 to 1.48)	1.00 (0.78 to 1.28)	1.15 (0.90 to 1.47)	1.19 (0.94 to 1.50)	1.00 (0.77 to 1.27)	1.13 (0.88 to 1.45)
NHL	1.14 (0.96 to 1.35)	0.84 (0.69 to 1.03)	1.15 (0.95 to 1.38)	1.11 (0.9 to 1.32)	0.84 (0.69 to 1.03)	1.11 (0.92 to 1.34)
Leukemia	1.01 (0.78 to 1.31)	0.95 (0.72 to 1.24)	1.36 (1.07 to 1.73)	1.01 (0.78 to1.33)	0.99 (0.75 to 1.30)	1.37 (1.07 to 1.76)

Analyses were adjusted by sex, year of birth, locality type, origin, cognitive score, and NOx.
HR=Hazard Ratio, CI= confidence interval; CNS = Central Nervous System, NHL = Non Hodgkin Lymphoma.

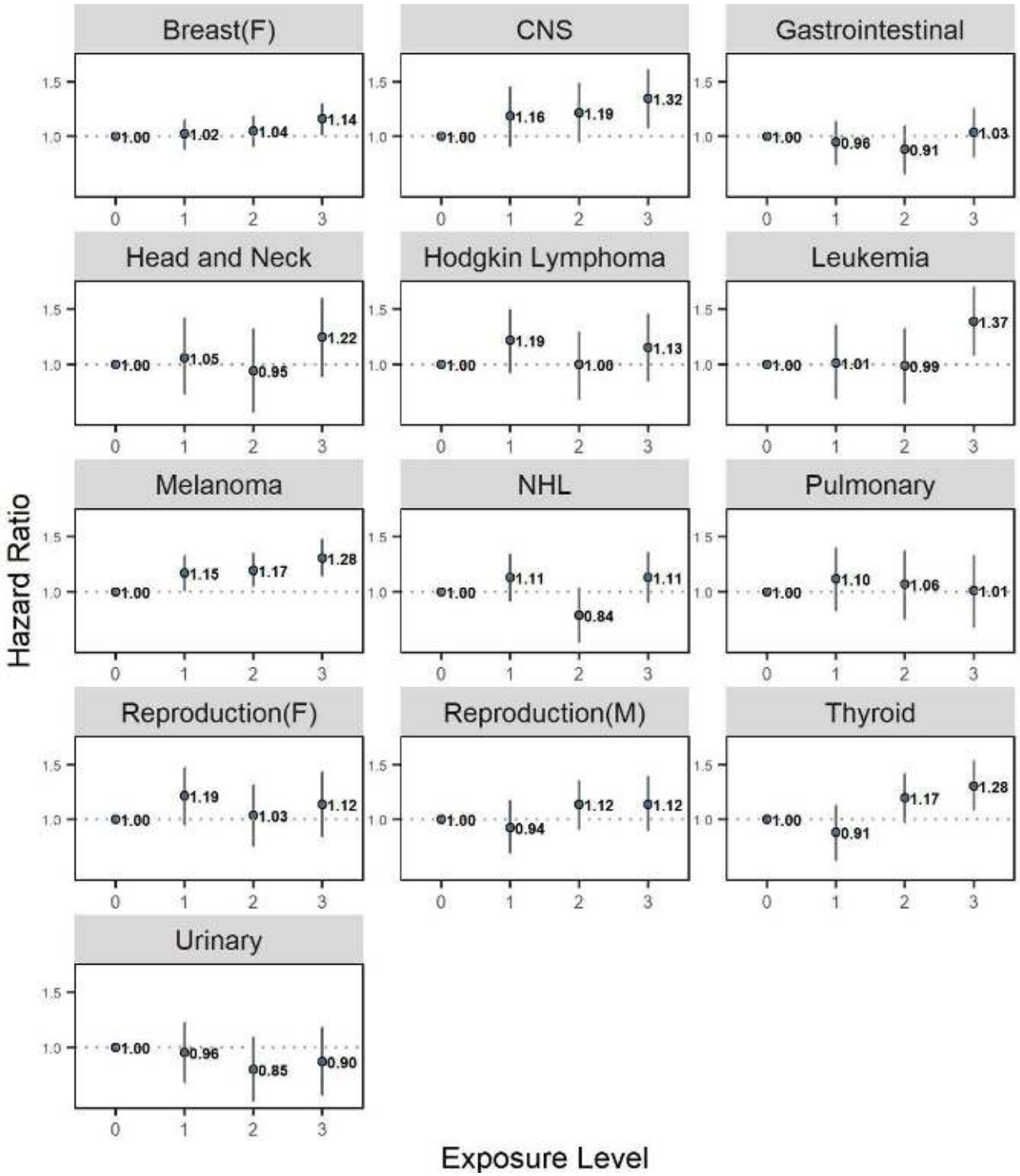

Figure 37. Associations of exposure to HBA-IAP with specific cancers, modeled using Cox proportional hazard regression, adjusted for sex, year of birth, locality type, origin, cognitive score and NOx. Sex was not included as a confounder in the models for female breast cancer, female/male reproduction. *N*=2,311,240, N cases: head and neck 1231; gastrointestinal 3909; pulmonary 1821; melanoma 4994; female breast 8315; female reproduction 1952; male reproduction 2830; urinary 2207; CNS 2040; thyroid 2799; Hodgkin lymphoma 1792; non-Hodgkin lymphoma 3152; lymphoid leukemia 1497. 0 = reference category (non-HBA residents), 1 = low HBA-IAP exposure level, 2= intermediate exposure, 3= high exposure.

To mitigate the possibility of confounding by coexisting illness, we conducted a sensitivity analysis that was limited to those with unimpaired health in adolescence (N=1,755,941). Our findings persisted. Moreover, the risk accentuated for male reproductive cancers, HR=1.19 (95% CI =0.97 to 1.46) in the unimpaired compared to HR=1.12 (95% CI=0.92 to 1.37) in the general cohort, for the highest exposure category. (See elaboration in Table 12 and Figure 39).

Table 12. Adjusted associations between HBA-IAP and cancer in subjects with unimpaired health at adolescence.

HBA-IAP	Low	Intermediate	High
Any Cancer	1.08 (1.02 to 1.13)	1.06 (1.01 to 1.12)	1.15 (1.09 to 1.21)
Head and Neck	1.10 (0.82 to 1.48)	0.89 (0.64 to 1.25)	1.19 (0.87 to 1.63)
Gastrointestinal	0.95 (0.80 to 1.14)	0.96 (0.81 to 1.15)	1.05 (0.87 to 1.26)
Pulmonary	1.09 (0.85 to 1.39)	1.08 (0.84 to 1.40)	1.04 (0.79 to 1.37)
Melanoma	1.12 (0.97 to 1.30)	1.16 (1.02 to 1.32)	1.30 (1.13 to 1.49)
Breast (Female)	1.02 (0.9 1 to 1.14)	1.02 (0.91 to 1.14)	1.11 (0.99 to 1.25)
Reproduction (Female)	1.12 (0.89 to 1.40)	1.06 (0.84 to 1.34)	1.10 (0.85 to 1.41)
Reproduction (Male)	0.93 (0.75 to 1.14)	1.12 (0.93 to 1.36)	1.19 (0.97 to 1.46)
Urinary	0.94 0.74 to 1.19)	0.81 (0.62 to 1.05)	0.88 (0.67 to 1.16)
CNS	1.16 (0.92 to 1.46)	1.04 (0.81 to 1.33)	1.32 (1.05 to 1.67)
Thyroid	0.91 (0.74 to 1.11)	1.17 (0.98 to 1.40)	1.28 (1.07 to 1.54)
Hodgkin Lymphoma	1.24 (0.97 to 1.59)	0.96 (0.73 to 1.26)	1.06 (0.81 to 1.40)
NHL	1.13 (0.94 to 1.36)	0.82 (0.66 to 1.01)	1.05 (0.86 to 1.29)
Leukemia	1.02 (0.77 to 1.36)	1.01 (0.76 to 1.36)	1.35 (1.03 to 1.76)

Analyses were adjusted by sex, year of birth, locality type, origin and cognitive score. *HR=Hazard Ratio, CI= confidence interval;* CNS = Central Nervous System, NHL = Non Hodgkin Lymphoma.

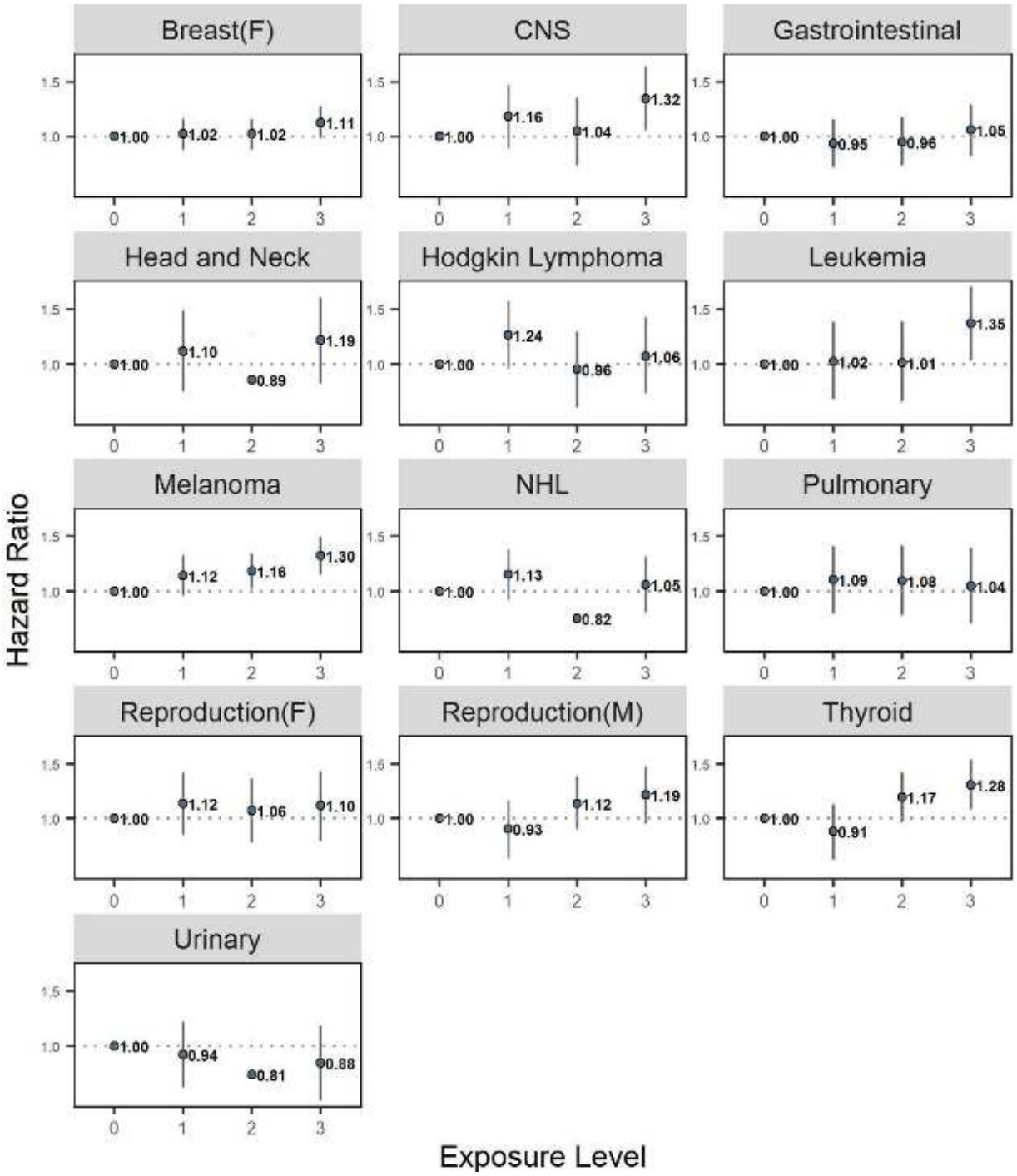

Figure 39. Associations of exposure to HBA-IAP with specific cancers, modeled using Cox proportional hazard regression, adjusted for sex, year of birth, locality type, origin, cognitive score and NOx, in participants with unimpaired health at adolescence. Sex was not included as a confounder in the models for female breast cancer, female/male reproduction. *N*=2,311,240, N cases: head and neck 1088; gastrointestinal 3443; pulmonary 1587; melanoma 4466; female breast 7758; female reproduction 1785; male reproduction 2640; urinary 1932; CNS 1738; thyroid 2799; Hodgkin lymphoma 1515; non-Hodgkin lymphoma 2772; lymphoid leukemia 1303. 0 = reference category (non-HBA residents), 1 = low HBA-IAP exposure level, 2= intermediate exposure, 3= high exposure.

Discussion

Main findings

The first part of this project is a cross-sectional study of health outcomes in adolescents at age 17 and their associations with HBA-IAP. We have specifically addressed health conditions that may be impacted by air pollution, focusing on asthma– a highly prevalent disease of the respiratory and immunologic systems.

Our results confirm that asthma prevalence is much higher in Israeli-born HBA adolescents than in their non-HBA counterparts. However, among HBA residents, the association between exposure to HBA-IAP (based on our exposure model) and either the crude prevalence of asthma or adjusted association with the disorder is not compliant with what is expected from causal relationships. Specifically, we think that the fact that the highest exposure group has the lowest prevalence and that this picture does not change even after adjustment for potential confounders argue against a causal relationship between childhood exposure to HBA-IAP and prevalent asthma during adolescence. Moreover, in most models, the highest risk of asthma is seen in the two lowest HBA-IAP categories. This is another piece of evidence implying that the higher asthma prevalence proportions in HBA may be caused by factors other than HBA-IAP.

Many possible exposure-response curves may be compliant with a causal relationship. For example, there may be a threshold level for the mean exposure to have an effect, or there may be a ceiling effect in which additional exposure beyond a certain level will not pose any additional risk. However, the non-monotonic exposure-response curves we have observed here are not compliant with any of these. Examinations of crude and adjusted associations of HBA-IAP with other atopic diseases and metabolic conditions do not change these general conclusions, as they do not provide support for causal relationships.

The second part of this project is a historical cohort study of associations between childhood and adolescence exposure to HBA-IAP with adult-onset cancer. Here, we found a positive monotonic exposure-response association between the exposure and any cancer in our crude analysis. The association was somewhat attenuated but remained clear and statistically significant with adjustment to potential confounders, with an HR of 1.16 (95% CI: 1.10-1.21) for the highest exposure group compared to non-HBA residents. Despite presenting more considerable uncertainty in the estimates (as expected), analyses stratified by the decade of birth or SES and restrictions to subjects with unimpaired health at baseline did not suggest that residual confounding is the cause for these findings. A separate inspection of cancer categories revealed positive monotonic associations in 6 out of 13 categories (female breast cancer, CNS, head and neck, leukemia, melanoma, and thyroid cancer) but inconsistent or null findings in the other categories (gastrointestinal organs, Hodgkin's lymphoma, NHL, reproduction organs (in males and females), urinary organs, and pulmonary cancer).

Possible explanations

The main findings we described above for asthma may have various explanations. Firstly, one should consider the simplest explanation: childhood exposure to HBA-IAP does not increase

the risk of asthma during childhood or adolescence. Therefore, we do not observe an exposure-response curve compliant with such a causal relationship. Indeed, because of the lack of measurements, we do not have a good characterization of the chemical and physical contents of the pollution emitted from the HBA industrial area. As explained in the background of this report, there are reasons to believe that this industry is emitting VOCs and heavy metals in different forms. However, we do not know what these are and what physical and chemical transformations and reactions these substances go through in the atmosphere. In addition, the most robust evidence in the literature about air pollution and asthma is with criteria pollutants and concerning asthma exacerbations. The evidence about onset of asthma, i.e., the development of asthma in a child without any history of the disease, is weaker and is also limited to criteria pollutants. Therefore, a lack of causal relationship between HBA-AIP and the risk of prevalent asthma in adolescence does not contradict previous epidemiological reports. Moreover, these findings do not imply a lack of causal relationship between HBA-IAP and asthma exacerbation and other diseases that were not included in this report.

Despite the simplest explanation stated above, other possible explanations exist as well. We acknowledge the possibility of residual (or unmeasured) confounding that may have prevented us from observing the real exposure-response curve. According to this possibility, as a group, the highest HBA-IAP exposure group is characterized by a high proportion of factors that protect it from asthma compared to all other exposure groups. These unmeasured protective factors cannot include factors we have adjusted for or stratified by. Therefore, it cannot result from differences among the groups that are strongly related to time trends, SES, populations group (as depicted by the school orientation variable), parents' origin, traffic-related air pollution, or $PM_{2.5}$. In addition, these protective factors must be prevalent in the highest exposure group and have a substantial impact on the risk of the relevant diseases (mainly asthma).

One potential confounder about which data were unavailable to us is smoking, including environmental (passive, second-hand) tobacco smoking, a strong risk factor for asthma. Smoking is also associated with factors such as SES that affect the spatial distribution of our study participants. Therefore, smoking is a potential unmeasured confounder that was not directly considered in our analyses, and very low levels of smoking in the highly exposed population, if they exist, may account for the findings. Unfortunately, we did not have individual-level or area-level data with relevant spatial resolution on smoking. However, one should note that our study included adolescents who were examined in a range of 50 years (1967 – 2017). While confounding structures are expected to show some changes over such an extended period, stratification by decades did not show any exposure-response curve that seems to support a causal effect in any of the periods. In addition, we did adjust our analyses for SES (at the small statistical area level) and school orientation – two variables that are expected to be associated with smoking prevalence, and reduce possible confounding by smoking. Therefore, we cannot completely rule out confounding by smoking, but several arguments do not support this possibility. Residual confounding from other variables, including genetic factors, is also possible. Genetic factors are partly controlled in stratified analyses by parental origin but cannot entirely be ruled out as potential confounders.

Another possible explanation for these exposure-response curves, which is compliant with a causal effect of HBA-IAP on prevalent asthma at age 17, is a possible interaction of this exposure with other (unmeasured) risk factors. For example, if the effect of HBA-IAP on the onset and persistence of asthma happens preferentially or exclusively when the person is co-exposed to a third factor, and this factor is not prevalent in the high HBA-IAP exposure areas, this may explain the low prevalence of asthma and other disorders in the highly exposed population. A possible factor that is known to interact with air pollutants in the causal mechanism of asthma onset is exposure to allergens. Indeed, there are plenty of possible allergens from different sources that may have such an interaction, and our study does not contain individual or area-level data on any of them. Therefore, it is possible that exposure to some allergens is causing the high prevalence of asthma and other atopic diseases that we observe in the low HBA-IAP exposure groups, with or without an interaction with HBA-IAP exposure. Indeed, this possible mechanism may also be compliant with a causal effect of HBA-IAP on the onset and persistence of asthma, if the relevant allergens are highly absent in the lowest HBA-IAP groups.

An additional possible explanation of our results that may be compatible with a causal relationship is information bias resulting from inaccurate residential addresses data. It is important to note that we have based our exposure assessment on addresses known at the medical evaluation, i.e., approximately at age 16-17. We used these addresses to calculate the mean exposure from birth to age 17. Some of the adolescents, possibly many, changed address between birth and age 17. In addition, none of them spent their entire life up to age 17 standing outside their home. They have naturally spent time indoors, in many other places, and for some of them – even their home addresses might have been different throughout most of the period. However, one should note that, as long as these error components do not distribute differently among adolescents with and without the relevant health condition we have examined, we can regard this error as non-differential (with respect to the outcome – for example, asthma). Such non-differential exposure error is expected, on average, to bias our associations towards the null, i.e., it is expected to yield associations in the same direction but weaker than expected. Since we have strong and statistically significant associations in our study that are not compatible with a causal exposure-response curve, a non-differential error of the exposure is unlikely to have caused these results.

On the other hand, we should not completely disregard the possibility of differential error in our exposure assessment due to a pattern of address changes that varies by the outcome. For example, suppose families with asthmatic children who lived in the highly exposed HBA area during their child's diagnosis are more likely to change their address before age 16 than families of asthmatic children from other regions. In that case, this would bias our crude and adjusted comparisons, which are based solely on prevalent asthma at age 17 and addresses around that age. Naturally, according to our model, most of the addresses in the highest HBA-IAP exposure are surrounding the industrial area or are downwind of it. It is certainly possible that parents of children who are diagnosed with asthma in this region think that it is better for their child to live in another area, and consequently move to a different area before age 17. In our analyses, these children will be counted in their new addresses, and will reduce the

observed prevalence of asthma at age 17 in the high HBA-IAP exposure group. Moreover, if these families move to less exposed HBA areas, they will increase the prevalence of asthma at age 17 in these areas. If such bias exists – it may explain our results and be compatible with a causal effect of HBA-IAP on asthma onset. However, we did not have data on addresses other than around age 17, so we could not assess the likelihood of this possibility.

Another possible explanation for our results that may comply with a causal relationship is that the exposure that matters for a causal effect is not the average exposure. For example, exposure to HBA-IAP varies by the time of the day. It is determined to a large extent by meteorological factors such as wind speed, wind direction, and atmospheric stability. Specifically, the spatial distribution of the nighttime IBA-IAP exposure during the night is different from that of the daytime. The concentrations of pollutants also vary by hour of the day and the presence of children and adolescents indoors and in their residential addresses. In this study, these factors were not considered in our epidemiological models, but some combination of these factors may be biasing our results. These issues require careful characterization and can be explored in future studies.

The associations with cancer diseases are more consistent with a causal effect of HBA-IAP. It is expected that such an effect will increase the risk of specific cancers but not others and that when such an increase is observed – it will follow a positive monotonic exposure-response curve. Notably, the list of cancer categories that presented clear associations with HBA-IAP contains several cancers related to endocrine systems (thyroid, breast, melanoma), suggesting involvement of endocrine disruptors in the exposure mix. However, both breast and thyroid cancers are often detected early, and melanoma is often associated with origin, leaving some room for possible residual confounding due to incomplete control of SES, screening services or origin factors.

It is also notable that some cancers that are typically associated with ambient air pollution – mainly pulmonary and prostate – are not associated with HBA-IAP. The lack of association with pulmonary and prostate cancers may be explained because these diseases are more common in older ages; most of the follow-up in our cohort occurred in young and middle-aged adults. On the other hand, the associations of leukemia and CNS cancers with HBA-IAP are consistent with prior epidemiological investigations of proximity to petrochemical complexes other than HBA and literature regarding the role of regarding the role of solvents and VOCs in the etiology of these tumors.[39]

The increased adjusted risk of “any cancer” may suggest that HBA-IAP exposure contains several chemical agents that act in various biological mechanisms to increase the risk of various cancer diseases. Our study cannot detect these agents or mechanisms as it lacks exposure assessment for specific chemicals. Still, it provides new evidence for the hypothesis that residential exposure to HBA-IAP increased the risk of cancer in our cohort and points to 6 out of 13 cancer categories in which the association is evident. Our exposure model enabled us to examine exposure-response curves, which provide more robust support for a causal hypothesis. The individual-level data we used allowed us to undertake a historical cohort study instead of the commonly-used ecological design that is much more prone to confounding. The sample size created statistical power to examine specific cancer categories

and carry out various stratified analyses – that further excluded potential residual confounding. Together, these structural components of the study permit more decisive conclusions than prior studies regarding the possibility of causal effects of HBA-IAP on the risk of cancer, despite the unavoidable observational nature of the study.

Main strengths and limitations

The main limitation of the study of health outcomes at age 17 is its cross-sectional design, limiting causal inference. This design is a result of the nature of the dataset we used, in which only one address and only one medical evaluation are given per participant, both around age 17. Ideally, one would need data on addresses from birth to diagnosis or end of follow-up, and relevant diagnosis date, to make a more robust inference on causal relations between exposure to pollution and risk of the onset of asthma. However, such data were not available to us. On the other hand, the current dataset is unique in that it was based on a mandatory medical screening. As such, it also includes a very large and comprehensive country-wide population-based sample that spans over 50 years of exposure and outcome data.

Another limitation is our exposure data, which is not ideal for an environmental epidemiology study. We have made a substantial effort to classify our study population by exposure to HBA-IAP. However, this exposure assessment lacks temporal variability, characterization of the relevant pollutants, and direct measures of these pollutants – components that would have improved it if they were available. Prior attempts to characterize HBA exposure were based on either distance from the industrial area, measurement of criteria pollutants, or simply comparing HBA to non-HBA residents.[43–46,48] We have used an innovative approach to categorize the HBA population by levels of exposure to HBA-IAP, and we argue that this approach is better than the other available measure or assessment method of historical HBA-IAP unmeasured pollutants. In addition, it enabled us to examine exposure-response relations, an important component for causal inference.

Another limitation is our lack of data on addresses before age 17 and lack of data about school addresses, which may account for a substantial portion of the exposure. We have limited our sample to Israeli-born adolescents to restrict this problem, but we acknowledge that this is one of the main limitations of this study. For the cancer study, we also lacked addresses after age 17 and thus could not differentiate between subjects for whom the exposure persisted and those for whom it ceased. In addition, we had no data about exposures after age 17, including occupational exposures, habits, family history of cancer, use of early detection tests, changes in BMI ets. However, epidemiological theory can predict the bias created by such inaccurate exposure data and the circumstances under which a causal relationship is hidden in our results, as discussed above for differential error of address change.

Our study is also limited by the outcomes we have examined. The most important limitation is that at age 17, we used prevalent health conditions without the ability to analyze disease onset, disease exacerbation, or further clinical information. On the other hand, we have examined several health conditions, including asthma, atopic diseases, metabolic and cardiovascular measures, and prospectively-assessed adult-onset cancer, covering several possible health effects of the pollutants.

Again, we also acknowledge our inability to better adjust for potential confounders, especially for smoking, including environmental tobacco smoking. In this context, we have examined possible residual confounding by various stratification methods, different exposure categorization methods, and proxies adjustments. The fact that these additional analyses did not materially change the exposure-response curve mitigates the possibility that our main findings result from residual confounding.

Conclusions

The analyses of health outcomes at age 17 did not support causal relationships between HBA-IAP and prevalent asthma, other atopic diseases, or major metabolic conditions at age 17. On the other hand, the finding of increased crude and adjusted risk of any cancer with increased HBA-IAP exposure strengthens the hypothesis that HBA-IAP increases cancer risk in the area. This conclusion has major public health implications, as the industrial zone in HBA is located in a highly-populated area.

References

1. Irigaray P, Newby JA, Clapp R, et al. Lifestyle-related factors and environmental agents causing cancer: An overview. *Biomed Pharmacother*. 2007;61(10):640-658. doi:10.1016/j.biopha.2007.10.006
2. Rumchev K, Brown H, Spickett J. Volatile organic compounds: Do they present a risk to our health? *Rev Environ Health*. 2007;22(1):39-55. doi:10.1515/REVEH.2007.22.1.39
3. Viegi G, Simoni M, Scognamiglio a, et al. Indoor air pollution and airway disease. *Int J Tuberc Lung Dis*. 2004;8(12):1401-1415. http://www.ncbi.nlm.nih.gov/pubmed/15636485
4. Miller RL, Peden DB. Environmental effects on immune responses in patients with atopy and asthma. *J Allergy Clin Immunol*. 2014;134(5):1001-1008. doi:10.1016/j.jaci.2014.07.064
5. Weisel CP. Assessing exposure to air toxics relative to asthma. *Environ Health Perspect*. 2002;110 Suppl(April):527-537. doi:10.1289/ehp.02110s4527
6. Delfino RJ. Epidemiologic evidence for asthma and exposure to air toxics: linkages between occupational, indoor, and community air pollution research. *Environ Health Perspect*. 2002;110 Suppl(October 2001):573-589. doi:10.1289/ehp.02110s4573
7. *Health Status in Israel 2010. Israeli Center for Disease Control (ICDC).*; 2011.
8. Clark NA, Demers PA, Karr CJ, et al. Effect of early life exposure to air pollution on development of childhood asthma. *Environ Health Perspect*. 2010;118(2):284-290. doi:10.1289/ehp.0900916
9. Jan AT, Azam M, Siddiqui K, Ali A, Choi I, Haq QMR. Heavy metals and human health: Mechanistic insight into toxicity and counter defense system of antioxidants. *Int J Mol Sci*. 2015;16(12):29592-29630. doi:10.3390/ijms161226183
10. Rosa MJ, Benedetti C, Peli M, et al. Association between personal exposure to ambient metals and respiratory disease in Italian adolescents: a cross-sectional study. *Bmc Pulm Med*. 2016;16:6. doi:10.1186/s12890-016-0173-9
11. Dunea D, Iordache S, Liu H-Y, Bøhler T, Pohoata A, Radulescu C. Quantifying the impact of PM2.5 and associated heavy metals on respiratory health of children near metallurgical facilities. *Environ Sci Pollut Res*. 2016;(13). doi:10.1007/s11356-016-6734-x
12. Shuster-Meiseles T, Shafer MM, Heo J, et al. ROS-generating/ARE-activating capacity of metals in roadway particulate matter deposited in urban environment. *Environ Res*. 2016;146:252-262. doi:10.1016/j.envres.2016.01.009
13. Yang CY, Wang J Der, Chan CC, Hwang JS, Chen PC. Respiratory symptoms of primary school children living in a petrochemical polluted area in Taiwan. *Pediatr Pulmonol*. 1998;25(5):299-303.
14. Chiang TY, Yuan TH, Shie RH, Chen CF, Chan CC. Increased incidence of allergic rhinitis, bronchitis and asthma, in children living near a petrochemical complex with SO2 pollution. *Environ Int*. 2016;96:1-7. doi:10.1016/J.ENVINT.2016.08.009
15. White CL, Pistell PJ, Purpera MN, et al. Effects of High Fat Diet on Morris Maze Performance, Oxidative Stress, and Inflammation in Rats: Contributions of Maternal Diet. *Neurobiol Dis*. 2009;35(1):3-13. doi:10.1016/j.nbd.2009.04.002.Effects
16. Wichmann FA, Muller A, Busi LE, et al. Increased asthma and respiratory symptoms in children exposed to petrochemical pollution. *J Allergy Clin Immunol*. 2009;123(3):632-638. doi:10.1016/j.jaci.2008.09.052

17. Rusconi F, Catelan D, Accetta G, et al. Asthma Symptoms, Lung Function, and Markers of Oxidative Stress and Inflammation in Children Exposed to Oil Refinery Pollution. *http://dx.doi.org/103109/02770903201053810 6*. 2011;48(1):84-90. doi:10.3109/02770903.2010.538106
18. De Moraes ACL, Ignotti E, Netto PA, Jacobson LDSV, Castro H, Hacon SDS. Wheezing in children and adolescents living next to a petrochemical plant in Rio Grande do Norte, Brazil. *J Pediatr (Rio J)*. 2010;86(4):337-344. doi:10.2223/JPED.2020
19. Yang CY, Wang J Der, Chan CC, Chen PC, Huang JS, Cheng MF. Respiratory and Irritant Health Effects of a Population Living in a Petrochemical-Polluted Area in Taiwan. *Environ Res*. 1997;74(2):145-149. doi:10.1006/ENRS.1997.3762
20. Kongtip P, Singkaew P, Yoosook W, Chantanakul S, Sujiratat D. Health effects of people living close to a petrochemical industrial estate in Thailand. *J Med Assoc Thai*. 2013;96 Suppl 5. Accessed March 29, 2022. https://pubmed.ncbi.nlm.nih.gov/24851575/
21. Tanyanont W, Vichit-Vadakan N. Exposure to volatile organic compounds and health risks among residents in an area affected by a petrochemical complex in Rayong, Thailand. *Southeast Asian J Trop Med Public Health*. 2012;43(1):201-211. Accessed March 29, 2022. https://pubmed.ncbi.nlm.nih.gov/23082571/
22. Yang CY, Cheng MF, Chiu JF, Tsai SS. Female Lung Cancer and Petrochemical Air Pollution in Taiwan. *https://doi.org/101080/00039899909602257*. 2010;54(3):180-185. doi:10.1080/00039899909602257
23. Weng H-H, Tsai S-S, Chiu H-F, Wu T-N, Yang C-Y. Association of Childhood Leukemia with Residential Exposure to Petrochemical Air Pollution in Taiwan. *Inhal Toxicol*. 2008;20(1):31-36. doi:10.1080/08958370701758734
24. García-Pérez J, López-Abente G, Gómez-Barroso D, et al. Childhood leukemia and residential proximity to industrial and urban sites. *Environ Res*. 2015;140:542-553. doi:10.1016/J.ENVRES.2015.05.014
25. Gottlieb MS, Shear CL, Seale DB. Lung cancer mortality and residential proximity to industry. *Environ Health Perspect*. 1982;45:157-164. doi:10.1289/EHP.8245157
26. López-Abente G, Fernández-Navarro P, Boldo E, Ramis R, García-Pérez J. Industrial pollution and pleural cancer mortality in Spain. *Sci Total Environ*. 2012;424:57-62. doi:10.1016/J.SCITOTENV.2012.02.047
27. Belli S, Benedetti M, Comba P, et al. Case-control study on cancer risk associated to residence in the neighbourhood of a petrochemical plant. *Eur J Epidemiol*. 2004;19(1):49-54. doi:10.1023/B:EJEP.0000013395.38701.C2
28. Yang CY, Chiu HF, Chiu JF, Kao WY, Tsai SS, Lan SJ. Cancer mortality and residence near petrochemical industries in taiwan. *J Toxicol Environ Health*. 1997;50(3):265-274. doi:10.1080/009841097160474
29. Pickle LW, Gottlieb MS. Pancreatic cancer mortality in Louisiana. *Am J Public Health*. 1980;70(3):256-259. doi:10.2105/AJPH.70.3.256
30. Pan BJ, Hong YJ, Chang GC, Wang MT, Cinkotai FF, Chin Ko Y. Excess cancer mortality among children and adolescents in residential districts polluted by petrochemical manufacturing plants in Taiwan. *http://dx.doi.org/101080/15287399409531908*. 2009;43(1):117-129. doi:10.1080/15287399409531908
31. Liu C-C, Chen C-C, Wu T-N, Yang C-Y. Association of brain cancer with residential exposure to petrochemical air pollution in Taiwan. *J Toxicol Environ Health A*.

2008;71(5):310-314. doi:10.1080/15287390701738491
32. Tsai SS, Tiao MM, Kuo HW, Wu TN, Yang CY. Association of bladder cancer with residential exposure to petrochemical air pollutant emissions in Taiwan. *J Toxicol Environ Health A*. 2009;72(2):53-59. doi:10.1080/15287390802476934
33. Ramis R, Diggle P, Boldo E, Garcia-Perez J, Fernandez-Navarro P, Lopez-Abente G. Analysis of matched geographical areas to study potential links between environmental exposure to oil refineries and non-Hodgkin lymphoma mortality in Spain. *Int J Health Geogr*. 2012;11. doi:10.1186/1476-072X-11-4
34. Wilkinson P, Thakrar B, Walls P, et al. Lymphohaematopoietic malignancy around all industrial complexes that include major oil refineries in Great Britain. *Occup Env Med*. 1999;56:577-580. doi:10.1136/oem.56.9.577
35. Di Salvo F, Meneghini E, Vieira V, et al. Spatial variation in mortality risk for hematological malignancies near a petrochemical refinery: A population-based case-control study. *Environ Res*. 2015;140:641-648. doi:10.1016/J.ENVRES.2015.05.022
36. Yuan T-H, Shen Y-C, Shie R-H, Hung S-H, Chen C-F, Chan C-C. Increased cancers among residents living in the neighborhood of a petrochemical complex: A 12-year retrospective cohort study. *Int J Hyg Environ Health*. 2018;221(2):308-314. doi:10.1016/j.ijheh.2017.12.004
37. Sans S, Elliott P, Kleinschmidt I, et al. Cancer incidence and mortality near the Baglan Bay petrochemical works, South Wales. *Occup Environ Med*. 1995;52:217-224. doi:10.1136/oem.52.4.217
38. Salerno C, Berchialla P, Palin LA, Vanhaecht K, Panella M. Cancer morbidity of residents living near an oil refinery plant in North-West Italy. *Int J Environ Health Res*. 2013;23(4):342-351. doi:10.1080/09603123.2012.733938
39. Domingo JL, Marquès M, Nadal M, Schuhmacher M. Health risks for the population living near petrochemical industrial complexes. 1. Cancer risks: A review of the scientific literature. *Environ Res*. 2020;186:109495. doi:10.1016/j.envres.2020.109495
40. Marquès M, Domingo JL, Nadal M, Schuhmacher M. Health risks for the population living near petrochemical industrial complexes. 2. Adverse health outcomes other than cancer. *Sci Total Environ*. 2020;730:139122. doi:10.1016/j.scitotenv.2020.139122
41. Ginsberg GM. Standardized mortality ratios for Israel, 1969-78. *Isr J Med Sci*. 1983;19(7):638-643.
42. Ginsberg GM. Standardized mortality ratios for Israel, 1983-86. *Isr J Med Sci*. 1992;28(12):868-877.
43. Paz S, Linn S, Portnov BA, Lazimi A, Futerman B, Barchana M. Non-Hodgkin Lymphoma (NHL) linkage with residence near heavy roads-A case study from Haifa Bay, Israel. *Heal Place*. 2009;15(2):636-641. doi:10.1016/j.healthplace.2008.10.004
44. Zusman M, Dubnov J, Barchana M, Portnov BA. Residential proximity to petroleum storage tanks and associated cancer risks: Double Kernel Density approach vs. zonal estimates. *Sci Total Environ*. 2012;441:265-276. doi:10.1016/j.scitotenv.2012.09.054
45. Eitan O, Yuval, Barchana M, et al. Spatial analysis of air pollution and cancer incidence rates in Haifa Bay, Israel. *Sci Total Env*. 2010;408(20):4429-4439. doi:10.1016/j.scitotenv.2010.06.031
46. Portnov BA, Dubnov J, Barchana M. Studying the association between air pollution and lung cancer incidence in a large metropolitan area using a kernel density function. *Socioecon Plann Sci*. 2009;43(3):141-150. doi:10.1016/j.seps.2008.09.001
47. Freedman D. The Ecological Fallacy. University of California. Published 2002. http://www.stat.berkeley.edu/~census/ecofall.txt

48. Rottenberg Y, Zick A, Barchana M, Peretz T. Organ specific cancer incidence in an industrial sub-district : a population-based study with 12 years follow up. *Am journals cancer Epidemiol Prev*. 2013;1(1):1-10.
49. Twig G, Yaniv G, Levine H, et al. Body-mass index in 2.3 million adolescents and cardiovascular death in adulthood. *N Engl J Med*. Published online 2016. doi:10.1056/NEJMoa1503840
50. Twig G, Reichman B, Afek A, et al. Severe obesity and cardio-metabolic comorbidities: a nationwide study of 2.8 million adolescents. *Int J Obes*. Published online 2019. doi:10.1038/s41366-018-0213-z
51. Furer A, Afek A, Sommer A, et al. Adolescent obesity and midlife cancer risk: a population-based cohort study of 2·3 million adolescents in Israel. *Lancet Diabetes Endocrinol*. Published online 2020. doi:10.1016/S2213-8587(20)30019-X
52. Twig G, Zucker I, Afek A, et al. Adolescent Obesity and Early-Onset Type 2 Diabetes. *Diabetes Care*. 2020;43(7):1487-1495. doi:10.2337/dc19-1988
53. Leiba A, Fishman B, Twig G, et al. Association of Adolescent Hypertension with Future End-stage Renal Disease. *JAMA Intern Med*. Published online 2019. doi:10.1001/jamainternmed.2018.7632
54. Leiba A, Twig G, Vivante A, et al. Prehypertension among 2.19 million adolescents and future risk for end-stage renal disease. *J Hypertens*. Published online 2017. doi:10.1097/HJH.0000000000001295
55. Moore E, Silverman BG, Fishier Y, et al. An assessment of the completeness and timeliness of the Israel national cancer registry. *Isr Med Assoc J*. 2021;23(1):23-27. Accessed April 3, 2022. https://www.ima.org.il/MedicineIMAJ/viewarticle.aspx?year=2021&month=01&page=23
56. Yuval, Broday DM. High-resolution spatial patterns of long-term mean concentrations of air pollutants in Haifa Bay area. *Atmos Environ*. 2006;40(20):3653-3664. doi:10.1016/j.atmosenv.2006.03.037
57. Asaf D, Pedersen D, Peleg M, Matveev V, Luria M. Evaluation of background levels of air pollutants over Israel. *Atmos Environ*. Published online 2008. doi:10.1016/j.atmosenv.2008.08.011
58. Levy I, Levin N, Yuval Y, Schwartz JD, Kark JD. Back-extrapolating a land use regression model for estimating past exposures to traffic-related air pollution. *Environ Sci Technol*. Published online 2015. doi:10.1021/es505707e
59. Yuval, Levy I, Broday DM. Improving modeled air pollution concentration maps by residual interpolation. *Sci Total Environ*. 2017;598:780-788. doi:10.1016/j.scitotenv.2017.04.117
60. Kloog I, Sorek-Hamer M, Lyapustin A, et al. Estimating daily PM2.5 and PM10 across the complex geo-climate region of Israel using MAIAC satellite-based AOD data. *Atmos Environ*. 2015;122:409-416. doi:10.1016/j.atmosenv.2015.10.004
61. Shtein A, Karnieli A, Katra I, et al. Estimating daily and intra-daily PM10and PM2.5in Israel using a spatio-temporal hybrid modeling approach. *Atmos Environ*. 2018;191(July):142-152.
62. Lumley TT; T. Package 'survival.' *R Top Doc*. 2015;(128(10)):28-33.

Supplementary Tables and Figures

Table S1. Distribution of subjects by exposure to HBA-IAP and asthma

HBA-IAP category	N	Asthma Cases, %, CI
Non-HBA	2216927	130873, 5.9% (5.9-5.9)
1	56480	3852, 6.8% (6.6-7.0)
2	56637	3897, 6.9% (6.7-7.1)
3	56587	3928, 6.9% (6.7-7.2)
4	56612	3411, 6.0% (5.8-6.2)

Table S2. Distribution of subjects by exposure to $PM_{2.5}$ and asthma

$PM_{2.5}$ Exposure Quintile	N	Asthma Cases, %, CI
1st (<20.6 μg/m^3)	166485	9706, 5.8% (5.7-5.9)
2nd (20.6-21.4 μg/m^3)	156784	9926, 6.3% (6.2-6.5)
3rd (21.6-22.4 μg/m^3)	159436	11219, 7.0% (6.9-7.2)
4th (22.4-23.1 μg/m^3)	163516	13522, 8.3% (8.1-8.4)
5th (23.1-27.7 μg/m^3)	156016	13766, 8.8% (8.7-9.0)

Table S3. Distribution of subjects by exposure to NOx and asthma

NOx Exposure Quintile	N	Asthma Cases, %, CI
1st (<5.63 ppb)	466042	12865, 2.8% (2.7-2.8)
2nd (5.63-10.5 ppb)	467242	25735, 5.5% (5.4-5.6)
3rd (10.5-14.7 ppb)	469614	31661, 6.7% (6.7-6.8)
4th (14.7-19.1 ppb)	465901	34950, 7.5% (7.4-7.6)
5th (19.1-130 ppb)	463367	35796, 7.7% (7.6-7.8)

Table S4. Distribution of subjects by exposure to HBA-IAP and additional atopic conditions

HBA-IAP category	Rhinitis Cases, %, CI	Conjunctivitis Cases, %, CI	Dermatitis Cases, %, CI	Rhinitis + Asthma, Cases, %, CI
Non-HBA	130873 5.9% (5.9-5.9)	8884 0.4% (0.39-0.41)	10906 0.49% (0.48-0.5)	34088 1.54% (1.52-1.55)
1	3852 6.8% (6.6-7.0)	269 0.48% (0.42-0.53)	281 0.5% (0.44-0.55)	1128 2.0% (1.88-2.11)
2	3897 6.9% (6.7-7.1)	241 0.42% (0.37-0.48)	284 0.5% (0.44-0.56)	1090 1.92% (1.81-2.04)
3	3928 6.9% (6.7-7.2)	243 0.43% (0.37-0.48)	292 0.52% (0.46-0.57)	1142 2.02% (1.9-2.13)
4	3411 6.0% (5.8-6.2)	218 0.38% (0.33-0.44)	230 0.41% (0.35-0.46)	906 1.6% (1.5-1.7)

Table S5. Distribution of subjects by exposure to HBA-IAP and BMI category

HBA-IAP	N	Underweight Cases, %, CI	Normal Cases, %, CI	Overweight Cases, %, CI	Obese Cases, %, CI
Non-HBA	2154518	292938 13.6% (13.55-13.64)	1543243 71.63% (71.57-71.7)	244320 11.34% (11.3-11.38)	74017 3.43% (3.41-3.46)
1	55307	6858 12.4% (12.12-12.67)	39786 71.94% (71.56-72.3)	6647 12.02% (11.75-12.29)	2016 3.64% (3.49-3.8)
2	56637	6783 12.18% (11.91-12.46)	40604 72.95% (72.58-73.32)	6432 11.55% (11.29-11.82)	1841 3.31% (3.16-3.46)
3	56587	6945 12.48% (12.21-12.76)	40499 72.81% (72.44-73.18)	6358 11.43% (11.16-11.69)	1821 3.28% (3.12-3.42)
4	56612	7147 12.86% (12.58-13.14)	39746 71.52% (71.14-71.89)	6575 11.83% (11.56-12.00)	2106 3.79% (3.63-3.95)

Table S6. Distribution of subjects by exposure to HBA-IAP and Hypertension

HBA-IAP	N	Hypertension Cases, %, CI
Non-HBA	2216927	10901, 0.57% (0.56-0.58)
1	56480	314, 0.66% (0.59-0.73)
2	56637	295, 0.62% (0.55-0.68)
3	56587	290, 0.59% (0.53-0.65)
4	56612	303, 0.61% (0.51-0.63)

Table S7. Equal-participants (tertiles) HBA-IAP exposure.

HBA-IAP category	N
Reference (non-HBA)	1,945,991
1	68,587
2	69,149
3	67,871

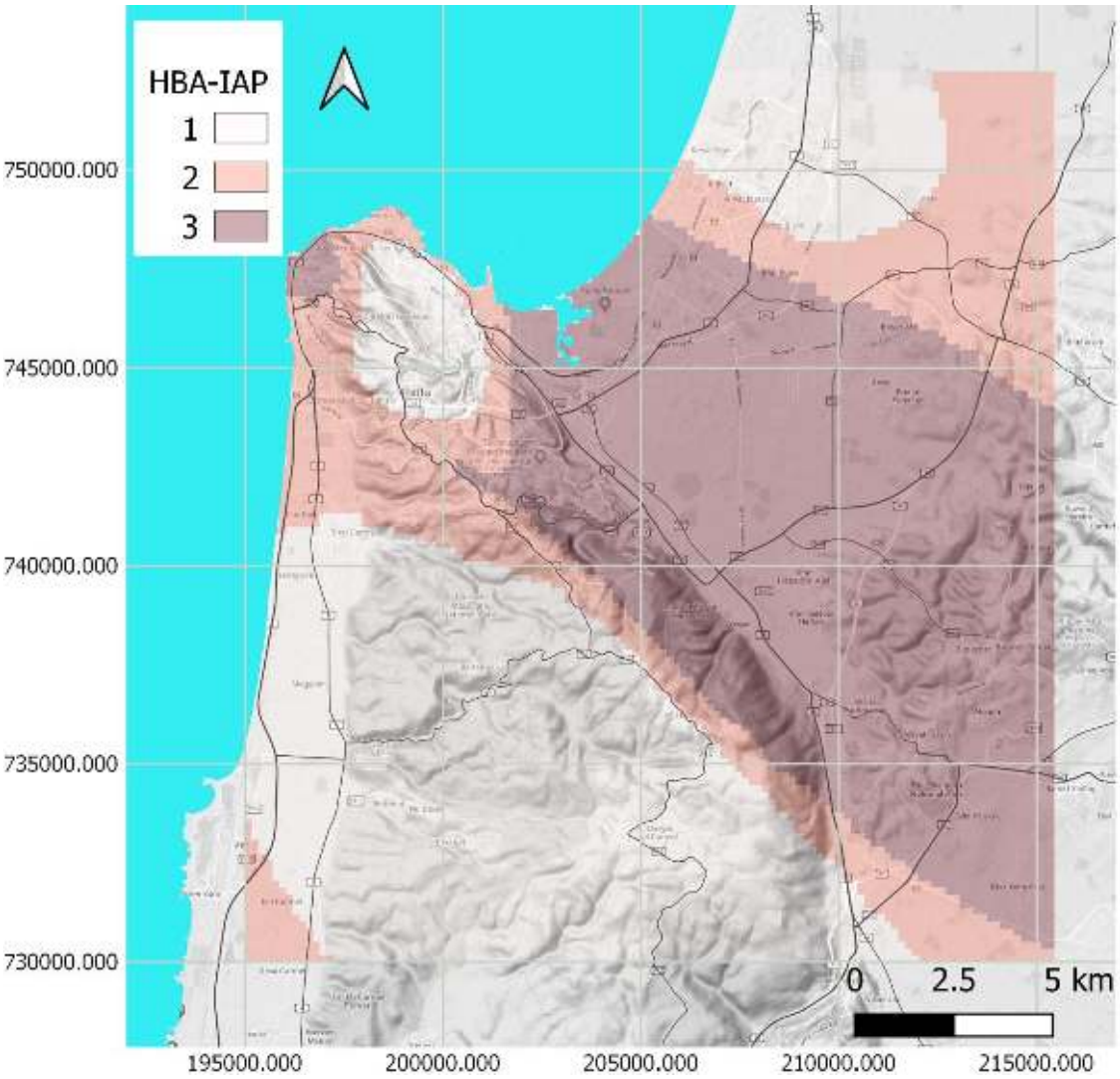

Figure S1. A map of the HBA-IAP exposure model,
with equal-participant number color-coded exposure categories 1-3.