Supplementary material for

Diagnostic validity of chronic kidney disease in health claims data over time: results from a cohort of community-dwelling older adults in Germany

List of contents

Table S1 – STARD 2015 Checklist

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	1, 3
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	3
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	4
	4	Study objectives and hypotheses	5
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	5
Participants	**6**	Eligibility criteria	5
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	5
	8	Where and when potentially eligible participants were identified (setting, location and dates)	5
	9	Whether participants formed a consecutive, random or convenience series	5
Test methods	**10a**	Index test, in sufficient detail to allow replication	6
	10b	Reference standard, in sufficient detail to allow replication	6
	11	Rationale for choosing the reference standard (if alternatives exist)	6
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	6-7
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	6-7
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	n/a
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	n/a
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	6-7
	15	How indeterminate index test or reference standard results were handled	5
	16	How missing data on the index test and reference standard were handled	5
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	6-8
	18	Intended sample size and how it was determined	5
RESULTS			
Participants	**19**	Flow of participants, using a diagram	8, Figure S1
	20	Baseline demographic and clinical characteristics of participants	8, Table 1
	21a	Distribution of severity of disease in those with the target condition	8, Table 1
	21b	Distribution of alternative diagnoses in those without the target condition	Figure 2
	22	Time interval and any clinical interventions between index test and reference standard	n/a
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	8-10, Table 2
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	8-10, Table 2
	25	Any adverse events from performing the index test or the reference standard	n/a
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	13
	27	Implications for practice, including the intended use and clinical role of the index test	10-13
OTHER INFORMATION			
	28	Registration number and name of registry	n/a
	29	Where the full study protocol can be accessed	n/a
	30	Sources of funding and other support; role of funders	15

Table S2 – Definition of chronic kidney disease (CKD) in BIS and claims data

BIS data		Health claims data
Reference group	eGFR based definition	Claims-based definition
CKD G3-5	eGFR <60 ml/min/1.73 m²	At least one of the following ICD-10-GM diagnoses (# displaying different operationalisations) within 6 months prior and 6 months post index date: #1 N18.3, N18.4, N18.5 #2 N18.3, N18.4, N18.5, N18.8x, N18.9, N19 #3 N18.1, N18.2, N18.3, N18.4, N18.5, N18.8x, N18.9, N19
CKD G3	eGFR 30–<60 ml/min/1.73 m²	At least one ICD-10-GM diagnosis N18.3 within 6 months prior to 6 months post index date
CKD G4-5	eGFR <30 ml/min/1.73 m²	At least one ICD-10-GM diagnosis N18.4 or N18.5 within 6 months prior to 6 months post index date
CKD G3-5$_{chronic}$	eGFR <60 ml/min/1.73 m² in two consecutive study visits; the latter study visit is defined as the index date	At least one of the following ICD-10-GM diagnoses (# displaying different operationalisations) within 6 months prior and 6 months post index date: #1 N18.3, N18.4, N18.5 #2 N18.3, N18.4, N18.5, N18.8x, N18.9, N19 #3 N18.1, N18.2, N18.3, N18.4, N18.5, N18.8x, N18.9, N19
CKD G3$_{chronic}$	eGFR 30–<60 ml/min/1.73 m² in two consecutive study visits; the latter study visit is defined as the index date	At least one ICD-10-GM diagnosis N18.3 within 6 months prior and 6 months post index date
CKD G4-5$_{chronic}$	eGFR <30 ml/min/1.73 m² in two consecutive study visits; the latter study visit is defined as the index date	At least one ICD-10-GM diagnosis N18.4 or N18.5 within 6 months prior and 6 months post index date

Abbreviations: BIS: Berlin Initiative Study; CKD: Chronic kidney disease; G3-5, G3, G4-5: Certain CKD stages as recommended in the Kidney Disease: Improving Global Outcomes (KDIGO) guidelines; eGFR: Estimated glomerular filtration rate; ICD-10-GM: International Statistical Classification of Diseases and Related Health Problems, 10th Revision, German modification.

Figure S1 – Flowchart of the BIS study. Shown are the number of participants seen at a respective study visit (with no-shows) from baseline to follow-up 4, periods in which data assessment took place, the number of excluded persons due to missing serum creatinine values, as well as the dropout cases due to death and loss to follow-up. Dropouts were considered as dead if a participant died before or within 3 months of the next intended follow-up (2 * x years after the baseline visit); all alive participants or later deaths were considered as lost to follow up with regard to the next pending follow-up. Participants who missed only single follow-ups in between (n=69) are displayed as "no-show".

Baseline: 2,069 participants
conducted from November 2009 until June 2011
Excluded because of missing serum creatinine values: 1 (0.05%)

Died: 148
Loss to follow up: 201

Follow-up 1: 1,699 participants (no-show: 21)
conducted from December 2011 until December 2013
Excluded because of missing serum creatinine values: 29 (1.7%)

Died: 152
Loss to follow up: 97

Follow-up 2: 1,440 participants (no-show: 31)
conducted from January 2014 until September 2015
Excluded because of missing serum creatinine values: 19 (1.3%)

Died: 172
Loss to follow up: 116

Follow-up 3: 1,166 participants (no-show: 17)
conducted from January 2016 until October 2017
Excluded because of missing serum creatinine values: 36 (3.1%)

Died: 136
Loss to follow up: 113

Follow-up 4: 934 participants
conducted from January 2018 until October 2019
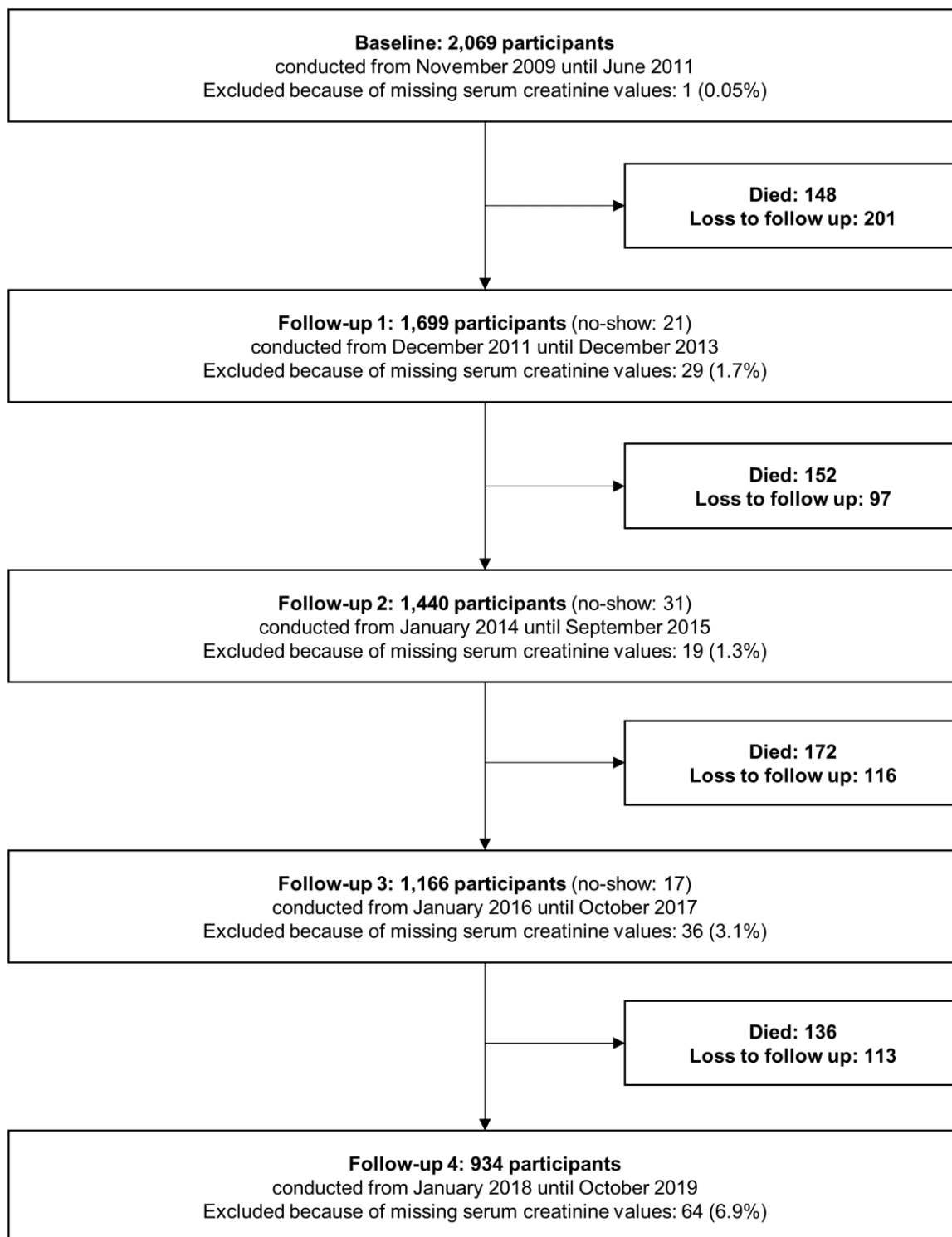Excluded because of missing serum creatinine values: 64 (6.9%)

Figure S2 – Sensitivity stratified by age and sex. Shades represent 95% confidence intervals and are interpolated between study visits for graphical display.
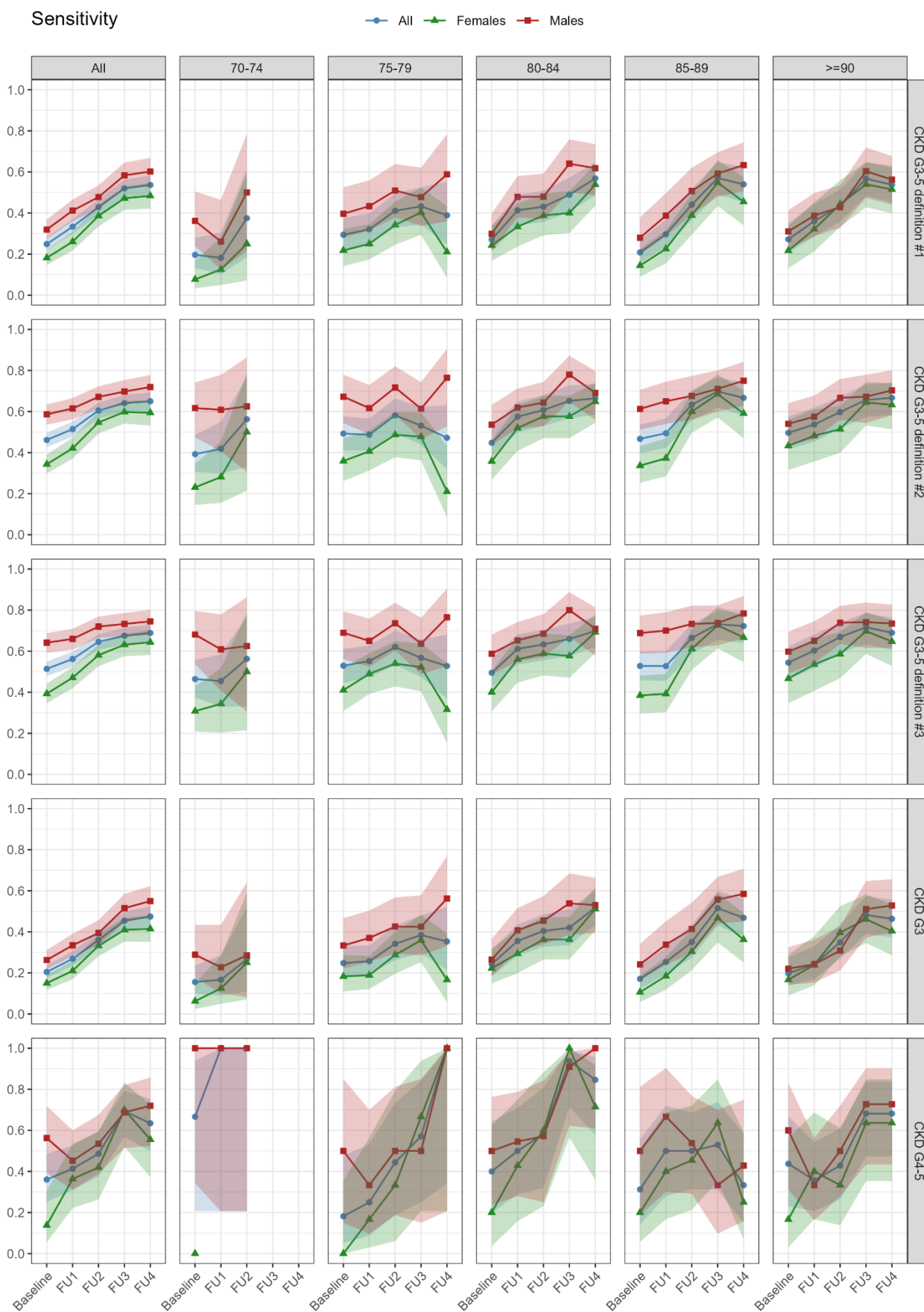
Figure S3 – Specificity stratified by age and sex. Shades represent 95% confidence intervals and are interpolated between study visits for graphical display.
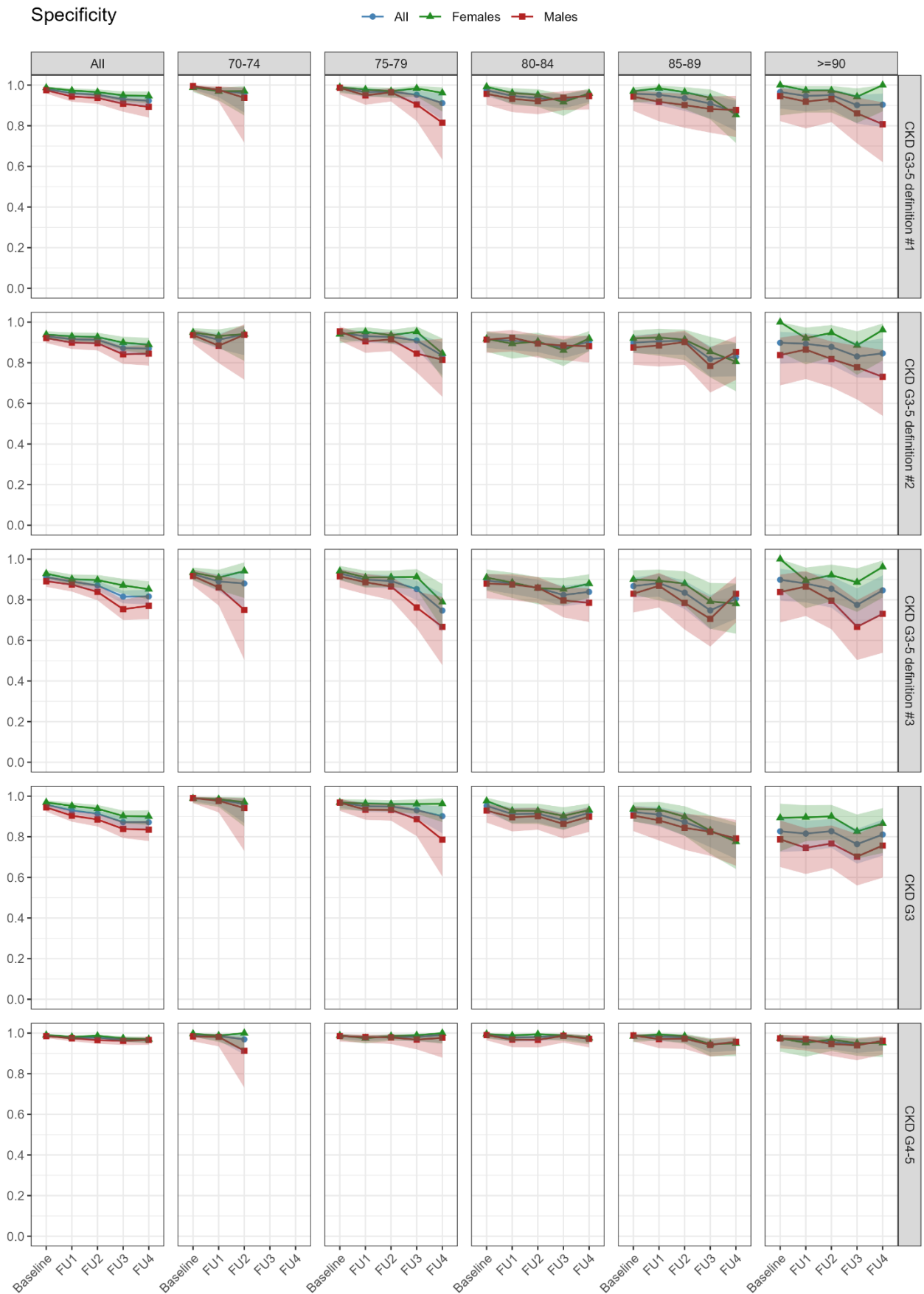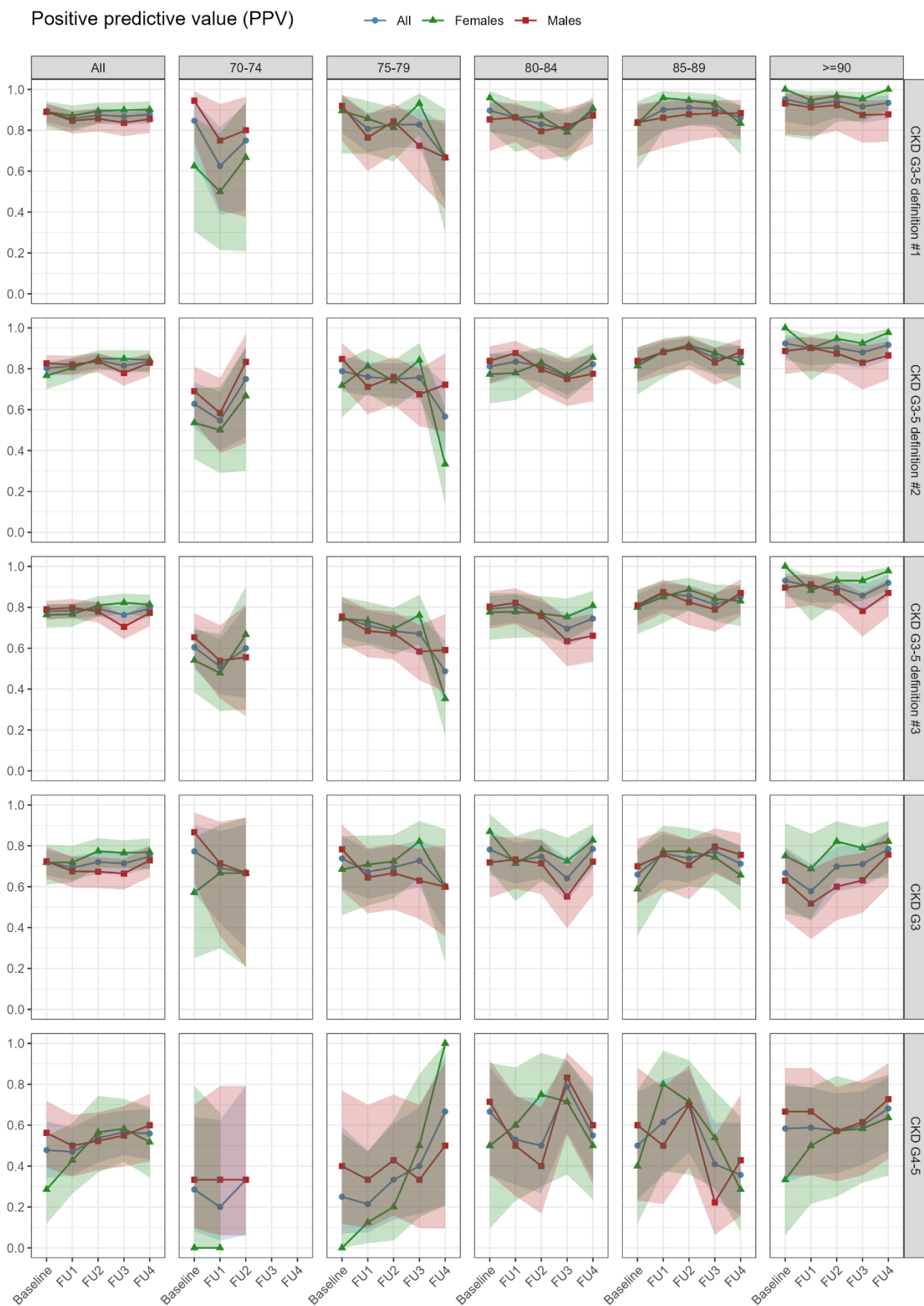
Figure S4 – Positive predictive values (PPV) stratified by age and sex. Shades represent 95% confidence intervals and are interpolated between study visits for graphical display.

Figure S5 – Negative predictive values (NPV) stratified by age and sex. Shades represent 95% confidence intervals and are interpolated between study visits for graphical display.
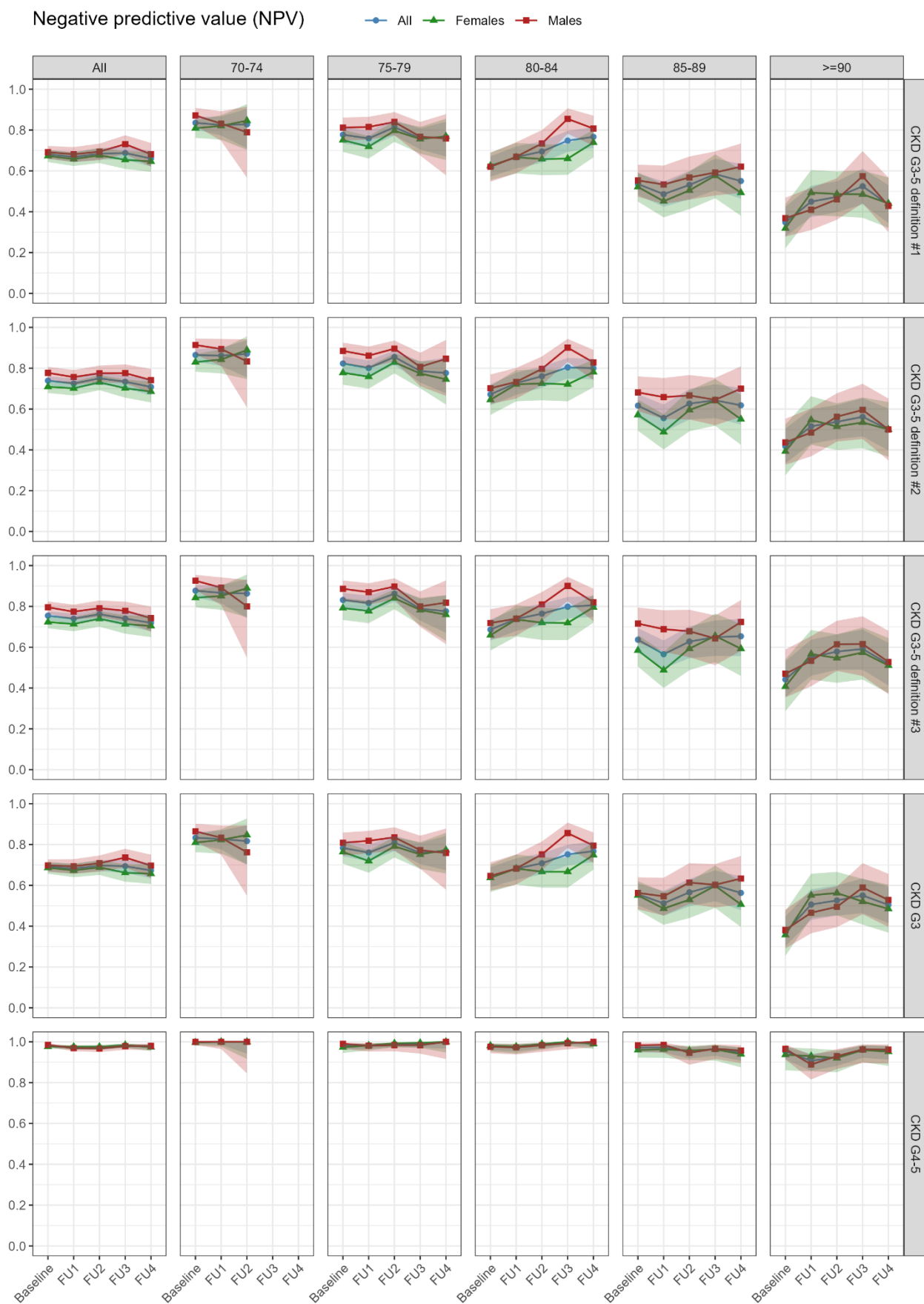
Figure S6 – Indicators of diagnostic validity (sensitivity, specificity, positive [PPV], and negative predictive values [NPV]) stratified by comorbidities for participants with either diabetes mellitus (without hypertension), arterial hypertension (without diabetes mellitus), both (diabetes mellitus and arterial hypertension), or none. Shades represent 95% confidence intervals and are interpolated between study visits for graphical display.
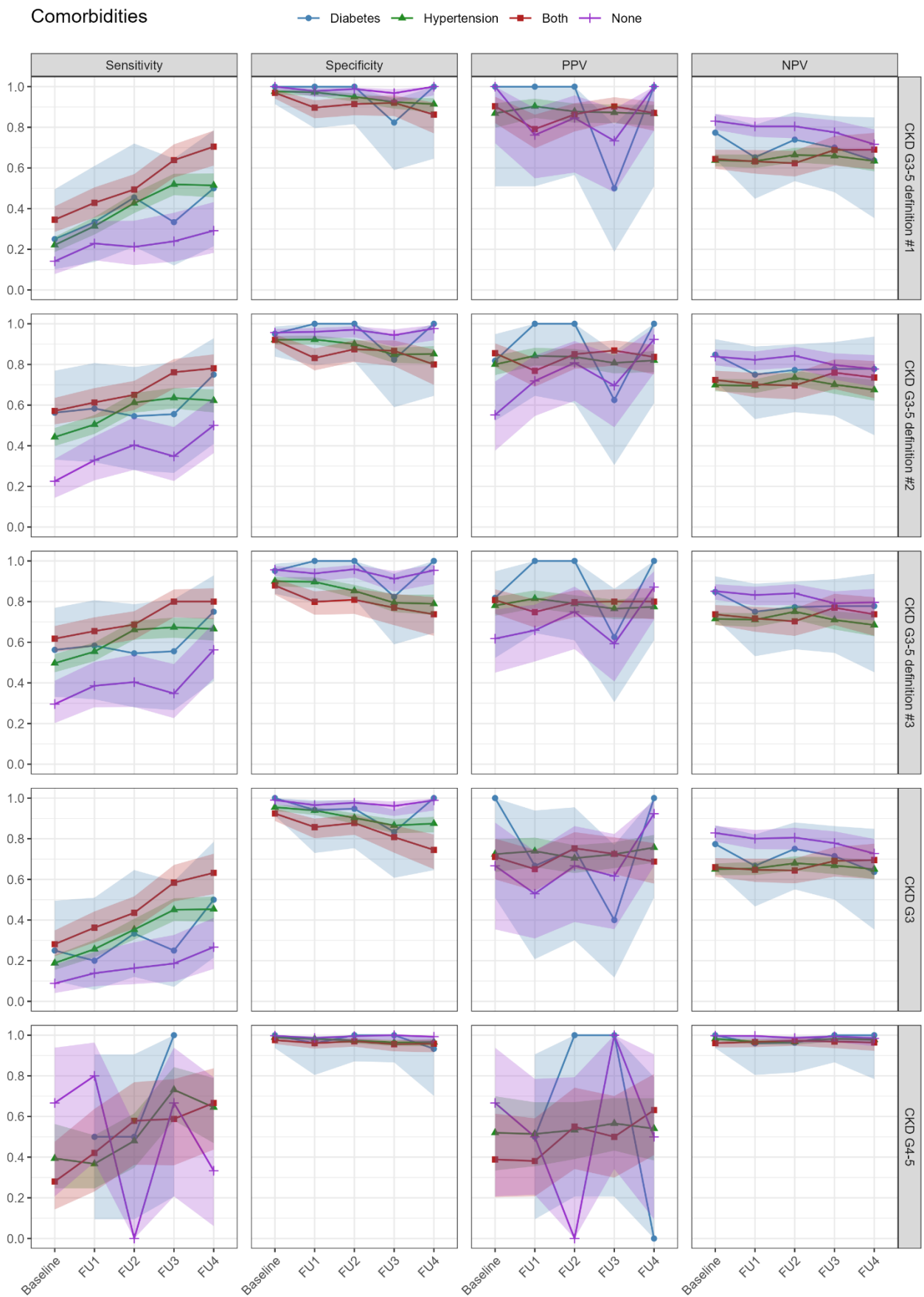
Figure S7 – Indicators of diagnostic validity (sensitivity, specificity, positive [PPV], and negative predictive values [NPV]) for both the single and chronic CKD definition (see Table S1 for the definitions). Shades represent 95% confidence intervals and are interpolated between study visits for graphical display.
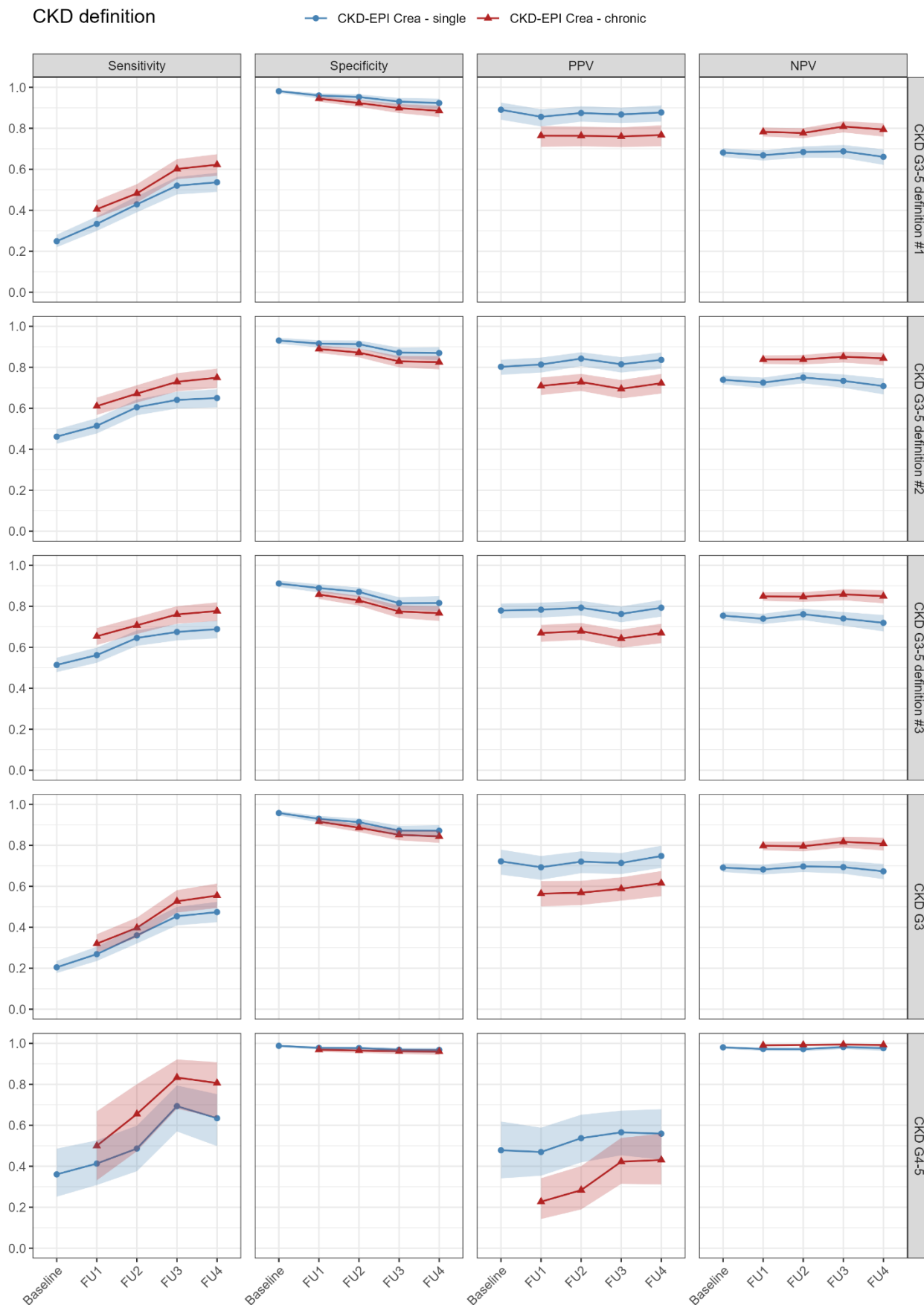
Figure S8 – Indicators of diagnostic validity (sensitivity, specificity, positive [PPV], and negative predictive values [NPV]) of different CKD definitions using weekly thresholds from one year preceding to one year following (preceding and following combined) a study visit as observation time for claims data diagnoses. Dashed lines represent the observation time window for the main analysis of +/- 6 months. Shades represent 95% confidence intervals and are interpolated for graphical display.