Kidney Medicine Original Research

Physicians' Recognition and Management of Kidney Disease: A Randomized Vignette Study Evaluating the Impact of the KDIGO 2012 CKD Classification System

Stein I. Hallan, Dena E. Rifkin, O. Alison Potok, Knut A. Langlo, Friedo W. Dekker, and Joachim H. Ix

Rationale & Objective: The Kidney Disease Outcome Quality Initiative (KDOQI) and Kidney Disease: Improving Global Outcomes (KDIGO) chronic kidney disease (CKD) classification systems published in 2002 and 2012, respectively, are recommended worldwide and based on strong epidemiologic data. However, their impact on CKD recognition and management is not well evaluated in clinical practice, and we therefore investigated whether they help physicians recognize and appropriately care for patients with CKD.

Study Design: Randomized vignette experiment with fractional factorial design based on 6 kidney-related scenarios and 3 laboratory presentation methods reflecting the CKD guidelines. Participants evaluated 1 of 3 subsets of the 18 vignettes (ie, 6 vignettes each with 4 answer alternatives).

Setting & Participants: 249 interns, general practitioners, and residents/fellows attending postgraduate meetings and courses in Norway and the United States.

Intervention: Kidney-related results (serum creatinine level and urinary albumin excretion) were presented as the "minimal data" (high/low levels), KDOQI-2002 (estimated glomerular filtration rate [eGFR] reported automatically), or KDIGO-2012 (eGFR + albuminuria categorization + risk for complications) laboratory report.

Outcome: CKD management choice by physicians.

Results: When kidney laboratory data were presented as the KDOQI-2002 report (automatic eGFR calculation), there was a significantly higher odds for correct patient management decisions compared with the minimal data report (OR, 1.57; $P < 0.001$). Additional significant improvement was obtained with the KDIGO-2012 report (OR, 2.28 for correct answer vs minimal data report [$P < 0.001$]; OR, 1.45 compared to KDOQI-2002 report [$P = 0.005$]). The KDIGO classification system improved physician management in 4 of the 6 clinical scenarios covering a wide range of kidney-related topics. Interaction analysis showed that general practitioners and those with 1 to 3 years of internal medicine experience had the greatest improvements with the new presentation techniques.

Limitations: Physicians' management was evaluated by theoretical scenarios rather than direct patient care.

Conclusions: Automatic GFR estimation, albuminuria categorization, and notification of the associated risk for complications improve most physicians' recognition and management of a wide range of CKD clinical scenarios.

Complete author and article information provided before references.

Correspondence to S.I. Hallan (stein.hallan@ntnu.no)

Kidney Med. 2(3):258-266. Published online February 22, 2020.

doi: 10.1016/j.xkme.2019.12.008

New clinical practice guidelines for the evaluation and management of chronic kidney disease (CKD) have recently been implemented worldwide.[1] However, their impact on patient management is not sufficiently tested, which is a deficiency of many disease classification systems. CKD has a very high prevalence and is associated with hypertension, cardiovascular disease, and mortality risk.[2] CKD management is often believed to be difficult for non-nephrology physicians. The Kidney Disease Outcome Quality Initiative (KDOQI) 2002 clinical practice guideline for CKD advocated automatic reporting of estimated glomerular filtration rate (eGFR), and the guideline relied strongly on eGFR for evaluation, classification, and risk stratification.[3] The Kidney Disease: Improving Global Outcomes (KDIGO) 2012 enhanced the KDIGO 2002 classification system by adding albuminuria levels to all patients and combining eGFR and albuminuria into a risk score for CKD-related complications.[4] The KDIGO 2012 CKD classification system has been shown to be strongly associated with prognosis across relevant subgroups in multiple studies among more than 1 million participants worldwide.[5-9] Expert panels recommend that the CKD classification system should be implemented into general practice and specialist care systems worldwide, specifically to enhance health care providers' recognition of CKD and improve related clinical treatment decisions.[10,11]

Editorial, p 235

However, despite representing a major change in public health policy, there is a critical lack of information about whether implementing the new guidelines improves patient management.[12-14] Clinical trials directly evaluating the utility of the KDIGO guideline do not appear feasible because double-blinding will not be possible, isolating physicians' competence and practice from structural effects is difficult, and such a trial may be considered unethical because the guideline is already widely implemented. Although not often

used in nephrology research,[15-17] a quantitative vignette study can overcome several of these barriers. It consists of an experiment in which clinical vignettes are administered in a randomized fashion and serve as the core element, and a traditional survey in parallel for collecting respondent-specific characteristics to be used as covariates in the analysis. The former increases internal validity, the latter improves external validity. Moreover, prior studies demonstrate that vignette studies accurately capture and reflect real-world behavior and decision making.[18-20]

We hypothesized that the 2012 KDIGO eGFR, albuminuria, and risk staging system would improve recognition and management of patients with CKD among physicians. In addition, we hypothesized that these staging systems would be useful over a wide range of clinical scenarios and would be particularly useful to non-nephrologist physicians and physicians with fewer years of clinical experience. We therefore designed and conducted a randomized clinical vignette study to test the impact of CKD laboratory reporting methods suggested by KDOQI and KDIGO over a wide range of clinical scenarios and simultaneously evaluate these findings relative to physician characteristics in a large and heterogeneous group of physicians in training and practice.

METHODS

We reviewed the KDIGO guideline to identify themes of importance for non-nephrologists.[1] Initially, a total of 12 clinical scenarios were created. Each scenario included information for kidney function (serum creatinine level) and urinary albumin excretion (urinary dipstick or albumin-creatinine ratio) with 4 different answer alternatives presented in a logical order (low to high risk or severity). These initial scenarios were subsequently reviewed and revised by consensus by a group of 7 experienced nephrologists. Three scenarios were discarded based on feedback, and the remaining 9 were tested in a pilot study among Norwegian general practitioners (n = 24) and hospital-based physicians at various levels of training (interns, residents, and fellows; n = 68) to evaluate their level of difficulty and ability to discriminate between high- and low-performing respondents. Three of the clinical scenarios were evaluated as too difficult (difficulty index < 0.3) or too easy (difficulty index > 0.8) and excluded.[21] The remaining 6 had a moderate difficulty index (0.48-0.67) and acceptable discrimination (discrimination index > 0.20); see Table S1 for additional details. These were therefore considered suitable and were carried forward. A full description of the clinical scenarios, answer alternatives, and the most correct answer is provided in Table S2.

Patient and Public Involvement

We invited interns, general practitioners/family medicine physicians (hereafter called GPs), and non-nephrology residents and fellows from Norway and the United States to participate during 2017 to 2018 (see Item S1 for details). Respondents were fully anonymous and gave information for only age, job type, and scientific and clinical practice experience. These physicians were randomly exposed to 6 different vignettes to evaluate the clinical utility of the KDIGO classification system. The study was evaluated by the Regional Ethical Committee and deemed exempt from review.

Vignette Experiment and Survey Design

A vignette is a brief carefully written description of a clinical situation designed to simulate key features of a real-world scenario.[20] Typically, a whole population of vignettes is constructed by systematically manipulating 1 or more factors across vignettes (experimental aspect) while other aspects are kept consistent (controlled aspects). We tested a wide range of clinical scenarios versus different laboratory presentation techniques, that is, a 6×3 factorial design for a total of 18 different vignettes.

First, each clinical scenario related to 1 or more of the main KDIGO guideline chapters (Table 1): (1) definition and classification of CKD (6 scenarios); (2) definition, identification, and prediction of CKD progression (2 scenarios); (3) management of progression and complications of CKD (3 scenarios); (4) other complications of CKD, including cardiovascular disease, medication dosage, patient safety, infections, hospitalizations, and caveats for investigating (4 scenarios); and (5) referral to specialists and models of care (2 scenarios).

Second, we used 3 different presentation techniques for kidney-related laboratory variables (see Table S3 for details and example): (1) the "minimal data" laboratory report indicated only whether the result was higher or lower than the reference range for creatinine level and urinary albumin-creatinine ratio, (2) the KDOQI-2002 laboratory report provided automatic calculating and reporting of eGFR consistent with the strong focus of KDOQI recommendations on kidney function level for classification of CKD presence and severity, and (3) the KDIGO-2012 laboratory report provided both eGFR and albumin

Table 1. Demographic Characteristics of Study Participants

	Interns (n = 93)	General Practitioners (n = 56)	Residents/ Fellows (n = 100)
Age, y	28.7 (5.7)	46.5 (10.2)	35.0 (4.1)
Job experience, total, y	1.0 (0.0)	13.9 (6.4)	5.6 (3.5)
Job experience, internal medicine, y	0.25 (0.0)	2.6 (1.4)	3.5 (1.3)
Authorship, yes	28 (30.1%)	7 (12.5%)	17 (17.0%)
PhD, yes	2 (2.2%)	1 (1.8%)	2 (2.0%)
Previous nephrology course, yes	10 (10.3%)	12 (21.4)	7 (7.0%)
Country, Norway	72 (77.4%)	56 (100.0%)	90 (90.0%)
Country, United States	21 (22.6%)	0 (0.0%)	10 (10.0%)

Note: N = 249. Data are given as mean (1 standard deviation) for continuous variables and number (percentage) for dichotomous variables.

categories and combined these into a risk for CKD-related complications (low, moderate, high, or very high risk) consistent with recommendations in the KDIGO guideline published in 2012.[1]

This vignette population was too large to be presented to each respondent, so a vignette subset was selected. Each respondent was provided one-third of all vignettes in a stratified way such that the level of each factor was evaluated at identical frequencies, that is, a mixed fractional factorial design. Accordingly, we created 3 different versions of the questionnaires (Table S4). The 3 different presentation techniques were rotated over the 6 clinical scenarios within each respondent, that is, each of the 3 methods were used in 2 vignettes for each of the respondents. The different questionnaires were randomly distributed to the participants, all with an intention to reduce selection bias due to inter-respondent differences in medical knowledge and experience.

Statistical Analyses

Analyses were performed using STATA, version 15 (Stata Corp). Background characteristics of participating physicians were described as mean with standard deviation and percentages. The proportion of correct answers was calculated for each laboratory presentation technique. Associations between correct questionnaire answers and laboratory data presentation methods, clinical scenarios, and physician characteristics were evaluated using logistic regression. Interactions between laboratory presentation technique and physician subgroups were tested on both the multiplicative and the additive scale.[22] Interaction on an additive scale means that the combined effect of 2 exposures is larger (or smaller) than the sum of the individual effects of the 2 exposures. Relative excess risk due to interaction (RERI) was calculated based on regression coefficients and adjusted for covariates, as needed.[23] Substituting odds ratios (ORs) from logistic regression for the relative risks normally used for calculating additive interaction will overestimate the relative risk when the baseline risk is very high, as in our study (proportion of correct answer expected to be 0.3-0.8).[24,25] We therefore used coefficients from log-binomial regression to avoid such bias.[25] $P < 0.05$ was considered statistically significant for all analyses, including interaction tests.

RESULTS

We randomly distributed 300 questionnaires to a heterogeneous group of physicians. The response rate in this vignette experiment evaluating whether the KDIGO classification improves physicians' recognition and management of patients with CKD was 83%. Background characteristics of the participating 93 interns, 56 GPs, and 100 hospital-based residents and fellows are shown in Table 1. Mean age was 35.2 years, length of postgraduate experience was 5.7 years, and internal medicine experience was 2.1 years. As expected, interns had the most limited level of experience, while GPs were older and had more experience.

The clinical scenarios are summarized in Table 2. Three different laboratory presentation techniques increasingly

Table 2. Brief Description of Clinical Scenarios (questions) Used to Evaluate Physicians' Management of CKD Patients

Scenario	Main KDIGO Theme		Clinical Scenario	Kidney Laboratory Data
1	2.1	Identification of CKD progression	35-y-old man with GN for 5 y, no biopsy, dipstick hematuria. lab data taken 2 y apart are shown: *How is his progression/prognosis?*	Scr, 1.41-1.67 mg/dL; UACR, 150-106 mg/g
2	3.1	Prevention of CKD progression	40-y-old woman has hypertension diagnosed. BP, 150/70 mm Hg. No diabetes and feels healthy: *Which BP medication and treatment goal?*	Scr, 0.96 mg/dL; UACR, 450 mg/g
3	4.5	Other complications; imaging studies, prevention of AKI	75-y-old man presenting with acute abdominal pain and some diarrhea. Furosemide, 2 0 mg, ×1 for mild heart failure. BP, 115/70 mm Hg: *Is it safe to order CT of the abdomen?*	Scr, 1.24 mg/dL; urine dipstick, A++
4	4.2	Other complications; interpretation of risk markers in CKD	58-y-old woman presenting with acute chest pain, now asymptomatic. Diabetes. Normal ECG and vital signs. TnT, 85 ng/L: *How is her risk for MI/ need for monitoring?*	Scr, 1.36 mg/dL; UACR, 1326 mg/g
5	4.4	Other complications; medication management, prevention of AKI	69-y-old man is planned for hip replacement. Well controlled hypertension treated with ACEi + Htz. He is otherwise healthy: *What precautions are needed to reduce risk for AKI?*	Scr, 1.30 mg/dL; urine dipstick, A+
6	5.1	Referral to specialist	55-y-old woman rescheduled due to kidney pathology at last lab tests 3 mo ago, now similar results. *What is her need for monitoring/referral to specialist?*	Scr, 2.85 mg/dL; urine dipstick, negative

Note: Scenarios describe laboratory test results that have been repeated and are considered representative. The scenarios/questions often relate to several KDIGO themes. See Table S2 for full description of scenario and answers to the multiple choice questions.
Abbreviations: ACEi, angiotensin-converting enzyme inhibitor; AKI, acute kidney injury; BP, blood pressure; CKD, chronic kidney disease; CT, computed tomography; ECG, electrocardiogram; GN, glomerulonephritis; Htz, hydrochlorothiazide; KDIGO, Kidney Disease: Improving Global Outcomes; lab, laboratory; MI, myocardial infarction; Scr, serum creatinine; TnT, troponin T; UACR, urinary albumin-creatinine ratio.

helped with the interpretation of kidney-related laboratory data as suggested by the new KDIGO classification system: the minimal data report with only high/low indicators served as the reference standard, the KDOQI-2002 report that automatically estimated GFR, and a KDIGO-2012–inspired laboratory report that automatically reported and combined eGFR and albuminuria categories into a risk for CKD-related complications. Each physician was provided 6 of the 18 clinical vignettes at random. We found no differences in response rates or physician background between the vignette sets (questionnaires) or between each individual vignette, consistent with the randomized allocation (Table S5).

The main effects of laboratory presentation technique, clinical scenario, job type, and internal medicine experience are shown in Figure 1A to D, respectively. On average, physicians chose the correct CKD management option in 47.9% of all cases when laboratory data were presented by the minimal data report. Corresponding results for the KDOQI-2002 and KDIGO-2012 laboratory reports were 59.2% and 67.7%, respectively. As expected, some clinical scenarios were more difficult than others, and interns and fellows and those with more internal medicine experience had higher proportions of correct responses. Table 3 shows results of formal testing using multivariable logistic regression. The unadjusted OR for providing the correct answer was 1.57 when kidney laboratory data were presented as eGFR (KDOQI-2002 report) as compared with indicating only high or low levels (minimal data report; $P < 0.001$). This corresponds to a relative risk of 1.25, that is, a 25% higher probability of providing the correct answer if the data were presented with automated eGFR reporting. Furthermore, when presenting laboratory results according to KDIGO recommendations (KDIGO-2012 report), the OR was 2.28 ($P < 0.001$) and equals relative risk of 1.40. The odds of providing the correct answer with the KDIGO-2012 report was also significantly higher when compared with the KDOQI-2002 report (OR, 1.45; $P = 0.005$). These findings were essentially unchanged with multivariable adjustment.

Next, we evaluated the effect of physicians' characteristics and experience. Age, scientific experience, total

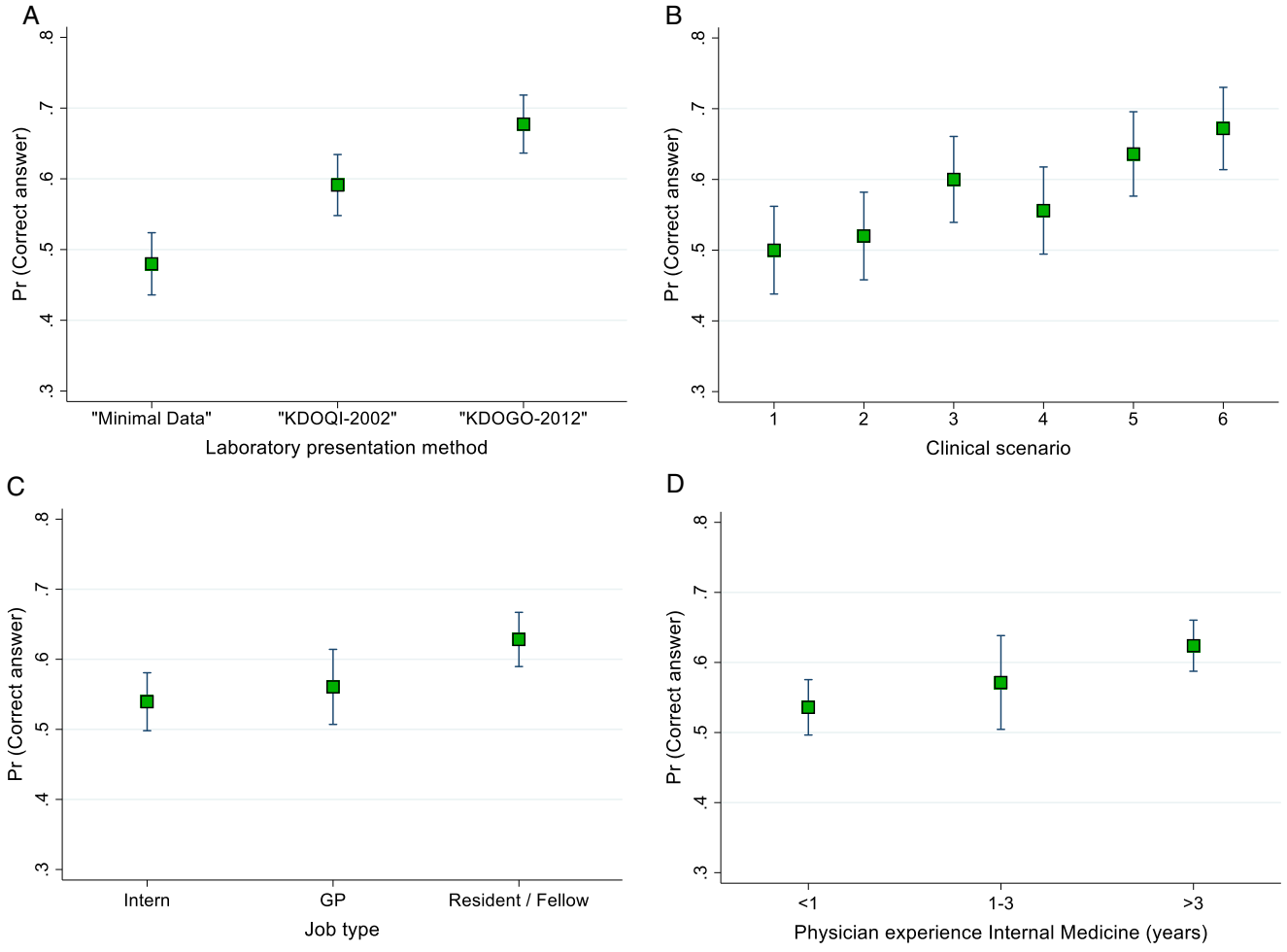

Figure 1. Main effects of experimental factors and major physician characteristics selected a priori: (A) laboratory presentation technique, (B) type of clinical scenario, (C) job type, and (D) internal medicine experience. Abbreviations: GP, general practitioner; KDIGO, Kidney Disease: Improving Global Outcomes; KDOQI, Kidney Disease Outcome Quality Initiative; Pr, probability.

Table 3. Association Between Variables and Probability of Correct Handling of Clinical Situation

	Unadjusted			Multivariable Adjustments		
Variable	**OR**	**95% CI**	***P***	**OR**	**95% CI**	***P***
Laboratory presentation method						
Minimal data	1.00	—	—	1.00	—	—
KDOQI-2002	1.57	1.22-2.01	<0.001	1.55	1.20-2.01	0.001
KDIGO-2012	2.28	1.76-2.94	<0.001	2.30	1.77-2.99	<0.001
Clinical scenario						
Question 1	1.00	—	—	1.00	—	—
Question 2	1.08	0.76-1.54	0.65	1.07	0.74-1.53	0.73
Question 3	1.50	1.05-2.14	0.03	1.50	1.04-2.16	0.03
Question 4	1.25	0.88-1.78	0.21	1.31	0.91-1.88	0.14
Question 5	1.74	1.22-2.50	0.002	1.77	1.23-2.56	0.002
Question 6	2.05	1.43-2.94	<0.001	2.17	1.50-3.15	<0.001
Job type						
Intern	1.00	—	—	1.00	—	—
GP	1.09	0.82-1.43	0.54	0.85	0.52-1.37	0.50
Resident/fellow	1.44	1.14-1.83	0.002	1.41	0.90-2.20	0.13
Age, per 10 y	1.02	0.91-1.14	0.76	0.94	0.76-1.16	0.58
Job experience, total, per 5 y	1.06	0.97-1.15	0.22	1.14	0.92-1.41	0.22
Internal medicine, per 1 y	1.10	1.04-1.16	0.002	0.99	0.85-1.17	0.95
Authorship, yes/no	0.89	0.69-1.44	0.37	1.05	0.78-1.40	0.88
PhD, yes/no	0.62	0.30-1.28	0.20	0.72	0.30-1.72	0.46
Country (US vs Norway)	0.74	0.55-1.01	0.06	0.70	0.47-1.05	0.09
Nephrology courses, yes/no	1.10	0.80-1.53	0.55	1.35	0.95-1.91	0.10

Note: N = 1,464 complete responses. Logistic regression analysis with correct/incorrect answer as dependent variable.
Abbreviations: CI, confidence interval; GP, general practitioner; KDIGO, Kidney Disease: Improving Global Outcomes; KDOQI, Kidney Disease Outcome Quality Initiative; OR, odds ratio.

length of career, country, and previous nephrology courses were not significantly associated with odds of providing the correct answer. However, residents and fellows had a higher probability of providing the correct answer as compared with interns (crude OR, 1.44; P = 0.002) or GPs (crude OR, 1.33; P = 0.04). Years of internal medicine experience were also associated with higher odds of providing the correct answer (OR, 1.10 per year of experience; P = 0.002).

Stratified analyses demonstrated that all physician subgroups were more likely to provide the correct response when provided with the KDOQI-2002 laboratory report compared with the minimal data report, and there was even greater improvement in correct responses when the KDIGO-2012 report was used for presenting kidney laboratory data (Fig 2A). We found a statistically significant multiplicative interaction (P = 0.04) when evaluating types of positions and the effect of the KDIGO-2012 report for the odds of providing the correct answer. Providing the KDIGO-2012 classification had greater improvement in correct responses among GPs than among interns, residents, or fellows. To illustrate on the additive scale, the combined effect of having the experience of a GP (vs an intern) and having kidney laboratory data presented with the KDIGO-2012 presentation technique (vs the minimal data) resulted in a 16–percentage point higher success rate than expected from the individual effects (RERI = 0.16 after correcting for internal medicine experience; P = 0.04). A similar finding was observed for years of internal medicine experience. Although the odds of providing the correct response improved with the KDOQI-2002 and KDIGO-2012 relative to the minimal data laboratory reporting across all subgroups of internal medicine experience, there was a statistically significant interaction whereby those with 1 to 3 years of internal medicine experience had the greatest improvement in providing the correct answer (P = 0.05). On the additive scale, those with 1 to 3 years of experience had a 17–percentage point higher success rate than expected when provided the KDIGO-2012 report as compared with a physician with less than 1 year of internal medicine using the minimal data report (RERI = 0.17 after adjusting for type of job position; P = 0.06; Fig 2B).

We tested the association between laboratory presentation method and the probability of correctly handling the 6 different clinical scenarios (Table 4). The KDIGO-2012 and KDOQI-2002 classification systems improved physicians' CKD management over a wide range of clinical scenarios. Kidney function decline assessment and specialist referral improved with the KDOQI-2002 laboratory report, and blood pressure management and acute kidney injury prevention improved significantly with the KDIGO-2012 as compared with the KDOQI-2002 report. We also evaluated the minimal data report in more details. In an ancillary study, physicians had a low ability to estimate GFR without any support, and only 43% of estimates

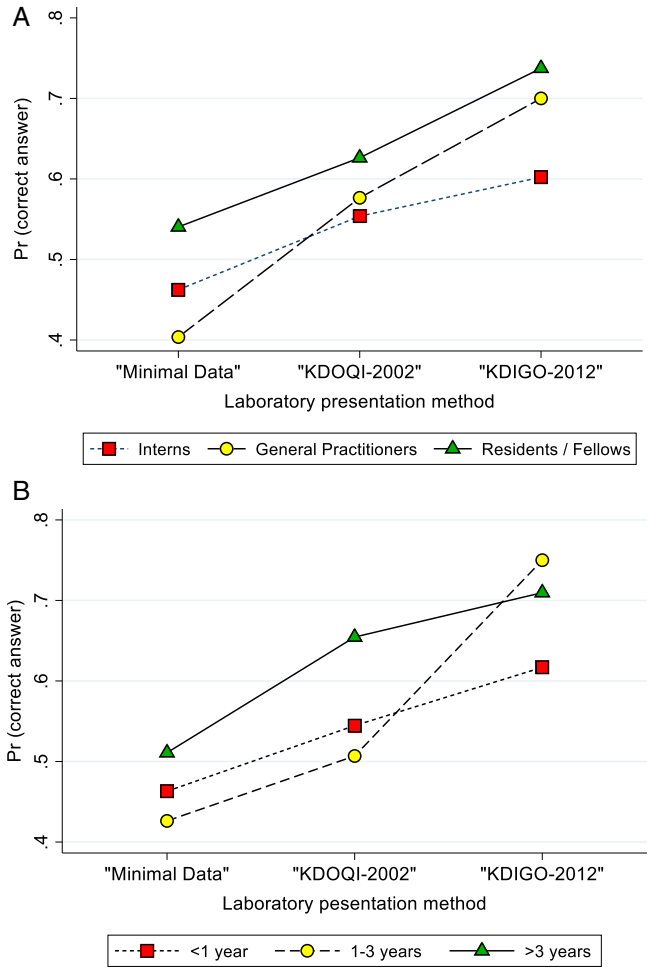

Figure 2. Interaction between laboratory presentation technique and (A) job type and (B) length of internal medicine experience, regarding the probability of correct CKD management. Abbreviations: GP, general practitioner; KDIGO, Kidney Disease: Improving Global Outcomes; KDOQI, Kidney Disease Outcomes Quality Initiative; Pr, probability.

were inside ±15% of the correct answer (proportion of physicians' responses within ±15% of the correct answer ranging from 25%-89% for 5 different cases based on various serum creatinine levels, age, and sex; Fig S1).

DISCUSSION

In this randomized vignette study, we demonstrate that the already widely implemented practice of automated laboratory calculation and reporting of eGFR when a serum creatinine level is reported significantly improved physicians' management decisions across a wide range of kidney-related scenarios. Adding albuminuria classification and a risk estimate of CKD-related complications as recommended by KDIGO in 2012 further improved physicians' performances. These reporting techniques improved management across all groups of physicians, but particularly among GPs and those with 1 to 3 years of experience in internal medicine.

Kidney disease is often considered to be a challenging part of medical practice for non-nephrologists,[26] and the concept of eGFR stages was conceptualized as an aid to assist GPs and non-nephrologist physicians to improve the diagnosis and management of patients. Thirty years ago, one of the few relevant studies on this topic demonstrated that physicians have considerable difficulty accurately estimating GFR based on serum creatinine level, sex, age, and weight.[13] We find that this is still the case among contemporary physicians. To our knowledge, all contemporary international kidney, hypertension, and cardiovascular disease guidelines now support the use of eGFR and albuminuria for CKD diagnosis and risk classification, and this foundation is strongly supported by large-scale epidemiologic studies demonstrating strong association with a variety of CKD-related outcomes.[1,5-9,27]

The implementation of automatic eGFR reporting as suggested by the KDOQI 2002 guideline has increased the recognition and referral of CKD as demonstrated by a meta-analysis summarizing reports published between 2002 and 2010.[28] However, a clinical benefit was more difficult to demonstrate. There was a slight increase in renin-angiotensin-aldosterone system–blocking drugs (0% to 6%) among patients with CKD with proteinuria, but no reduction in the use of nonsteroidal anti-inflammatory drug prescriptions and no improvement in correct dosing of antibiotics.[28] More recent publications on the KDOQI implementation have found reduced prescription of nonsteroidal anti-inflammatory drugs in patients with CKD[29] and a reduction in late referrals,[30] but no

Table 4. Association Between Laboratory Presentation Method and Probability of Correct Handling of Clinical Situation by Clinical Scenario

	KDOQI-2002 vs Minimal Data Laboratory Presentation			KDIGO-2012 vs KDOQI-2002 Laboratory Presentation		
Clinical scenario	**OR**	**95% CI**	***P***	**OR**	**95% CI**	***P***
Question 1 (How is his progression/prognosis?)	7.65	3.71-15.78	<0.001	1.51	0.77-2.95	0.22
Question 2 (Which BP medication and treatment goal?)	1.22	0.64-2.35	0.54	2.61	1.35-5.04	0.004
Question 3 (Is it safe to order CT of abdomen?)	1.16	0.61-2.23	0.64	0.64	0.32-1.12	0.11
Question 4 (How is her CV risk/need for monitoring?)	0.96	0.52-1.78	0.89	1.34	0.70-2.52	0.37
Question 5 (Which precautions are needed to avoid AKI?)	0.61	0.32-1.18	0.15	1.92	1.00-3.77	0.05
Question 6 (Is specialist referral/monitoring needed?)	2.47	1.31-4.70	0.006	2.85	1.28-6.36	0.01

Note: Multivariable logistic regression analysis on the effect of laboratory presentation method for each of the 6 different clinical scenarios.
Abbreviations: AKI, acute kidney injury; BP, blood pressure; CI, confidence interval; CT, computed tomography; CV, cardiovascular; KDIGO, Kidney Disease: Improving Global Outcomes; KDOQI, Kidney Disease Outcome Quality Initiative; OR, odds ratio.

improvement in classic quality indicators such as the proportions of patients starting hemodialysis treatment with arteriovenous fistulas or peritoneal dialysis was reported.[30]

The KDIGO CKD classification system, published in 2012, should intuitively be helpful to both GPs, hospital specialists, and nephrologists. To our knowledge, very few prior studies have formally evaluated the utility of this classification system. This is at odds with the introduction of new medications for which regulatory agencies typically require evidence of improvements in patient outcomes relative to contemporary standards of care.[31] A Canadian study reported shorter waiting lists for nephrology visits after integrating KDIGO 2012 recommendations and the derived Kidney Failure Risk Equation into the triage process, but long-term follow-up of patients triaged as low risk was not reported.[32] We found that automated eGFR reporting improved the odds for correct CKD classification and management by 50%. Adding interpretation of albuminuria plus a quantification of risk for complications improved performance by an additional 50%. The improvements were especially strong among GPs and physicians with 1 to 3 years of internal medicine experience. This finding is in accordance with general learning theory showing that learning curves typically follow a sigmoid function,[33] that is, with slowly accumulating small steps in the first phase (interns) when learners familiarize themselves with a topic and understand the basic definitions and terminology, followed by larger steps (GPs and residents), and then successively smaller ones later (fellows/nephrologists), as the learning reaches its limit.[34,35] We therefore believe the marked improvements demonstrated in our study should be clinically meaningful for individual patients and add important data to the scarce clinical evidence underlying the KDIGO classification system. Better identification and treatment of CKD, especially among those with high risk for complications and rapid progression, should help providers, policy makers, and health care systems target preventive measures and resources toward those most likely to benefit.[36-38] Further research is needed, and treatment initiatives should be teamed with field surveys and other efforts to study the clinical utility of the KDIGO classification system and the associated guidelines.

Strengths of our study include the randomized vignette design and evaluation in multiple settings and 2 countries. Our study also has important limitations. First, we only evaluated physicians' management of theoretical scenarios. However, a traditional randomized clinical trial is not feasible for the questions posed here and may be considered by some to be unethical.[20,39,40] Furthermore, physicians can change their behavior when directly observed or evaluated, whereas the vignette design creates a distance and is considered the most effective design for minimizing such effects.[20]

Second, although we exposed the physicians to clinical scenarios known to be important for CKD recognition and management and included physicians representing different specialties, levels of experience, and countries, the generalizability of our results remains untested in other settings and would benefit from confirmation. However, a vignette study is a hybrid of an experiment and a traditional survey and it therefore provides both internal and external validity.[41] Vignette studies are well suited to test specific questions about decision making, they have demonstrated generalizability to real-world situations, and they often overcome ethical and practical limitations of other designs.[18,19,42-44]

In conclusion, automatic GFR estimating substantially improves recognition and management of CKD. Adding information for categories of albuminuria and the associated risk for complications as recommended by KDIGO in 2012 further improves physicians' performance in recognizing and treating kidney disease. These strategies appear particularly useful for general practitioners and physicians with fewer years of internal medicine experience. Our data strongly support full implementation of the KDIGO 2012 CKD classification system into clinical practice.

SUPPLEMENTARY MATERIAL

Supplemental File (PDF)

Figure S1: Pilot study on unaided physician GFR estimates (mL/min/1.73 m^2) for 5 different patients (s-creatinine as μmol/L). Accuracy is calculated as the proportion of physicians` responses within ±15% of the correct answer (p15)

Item S1: Methods for vignette study and for ancillary study on unguided GFR estimation

Table S1: Level of difficulty and ability to discriminate high-scoring candidates from low-scoring candidates in a pilot study of 92 Norwegian physicians

Table S2: Full description of clinical scenarios (with KDIGO-2012 presentation technique for all laboratory results) with answer alternatives (most correct alternative marked in bold)

Table S3: Alternative presentation formats for kidney-relevant laboratory data

Table S4: Three different questionnaires (Q1-3) showing the rotation scheme of laboratory presentation techniques (minimal data, KDOQI-2002, and KDIGO-2012 laboratory reports) throughout the 6 clinical scenarios (1-6)

Table S5: Randomization of vignettes to the participants

ARTICLE INFORMATION

Authors' Full Names and Academic Degrees: Stein I. Hallan, MD, PhD, Dena E. Rifkin, MD, O. Alison Potok, MD, Knut A. Langlo, MD, Friedo W. Dekker, MD, PhD, and Joachim H. Ix, MD.

Authors' Affiliations: Department of Clinical and Molecular Medicine, Faculty of Medicine, Norwegian University of Science and Technology (SIH, KAL); Department of Nephrology, St. Olav Hospital, Trondheim, Norway (SIH, KAL); Division of Nephrology-Hypertension (DER, OAP, JHI), and Division of Preventive Medicine, Department of Family Medicine and Public Health (DER, JHI), University of California San Diego; Nephrology Section, Veterans Affairs San Diego Healthcare System (DER, OAP, JHI), San Diego, CA; and Department of Clinical Epidemiology, Leiden University Medical Center, Leiden, the Netherlands (FWD).

Address for Correspondence: Stein I. Hallan, MD, PhD, Department of Nephrology, St Olav Hospital, 7000 Trondheim, Norway. E-mail: stein.hallan@ntnu.no

Authors' Contributions: Study design: SIH; data collection: SIH, DER, OAP, KAL; data analysis: JHI, FWD, SIH; joint supervision/mentorship: JHI, FWD. Each author contributed important intellectual content during manuscript drafting or revision and accepts accountability for the overall work by ensuring that questions pertaining to the accuracy or integrity of any portion of the work are appropriately investigated and resolved.

Support: OAP is supported by the American Kidney Fund Clinical Scientist in Nephrology Fellow Program and Akebia Therapeutics, Inc. Funders did not have any role in study design; data collection, analysis, or reporting; or the decision to submit for publication.

Financial Disclosure: The authors declare that they have no relevant financial interests.

Acknowledgements: We thank the participants for their time and contribution to this study and the Unit of Applied Clinical Research at the Norwegian University of Science and Technology for help with optical reading of questionnaires.

Peer Review: Received September 26, 2019. Evaluated by 2 external peer reviewers, with direct editorial input from the Statistical Editor and the Editor-in-Chief. Accepted in revised form December 17, 2019.

REFERENCES

1. Levin A, Stevens PE. Summary of KDIGO 2012 CKD guideline: behind the scenes, need for guidance, and a framework for moving forward. *Kidney Int*. 2014;85(1):49-61.
2. Webster AC, Nagler EV, Morton RL, Masson P. Chronic kidney disease. *Lancet*. 2017;389(10075):1238-1252.
3. Levey AS, Coresh J, Balk E, et al. National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med*. 2003;139(2):137-147.
4. Stevens PE, Levin A; Kidney Disease: Improving Global Outcomes Chronic Kidney Disease Guideline Development Work Group Members. Evaluation and management of chronic kidney disease: synopsis of the Kidney Disease: Improving Global Outcomes 2012 clinical practice guideline. *Ann Intern Med*. 2013;158(11):825-830.
5. Hallan SI, Matsushita K, Sang Y, et al. Age and association of kidney measures with mortality and end-stage renal disease. *JAMA*. 2012;308(22):2349-2360.
6. Nitsch D, Grams M, Sang Y, et al. Associations of estimated glomerular filtration rate and albuminuria with mortality and renal failure by sex: a meta-analysis. *BMJ*. 2013;346:f324.
7. Wen CP, Matsushita K, Coresh J, et al. Relative risks of chronic kidney disease for mortality and end-stage renal disease across races are similar. *Kidney Int*. 2014;86(4):819-827.
8. Fox CS, Matsushita K, Woodward M, et al. Associations of kidney disease measures with mortality and end-stage renal disease in individuals with and without diabetes: a meta-analysis. *Lancet*. 2012;380(9854):1662-1673.
9. Matsushita K, Ballew SH, Astor BC, et al. Cohort profile: the Chronic Kidney Disease Prognosis Consortium. *Int J Epidemiol*. 2013;42(6):1660-1668.
10. National Clinical Guideline Centre (UK). *Chronic Kidney Disease (Partial Update): Early Identification and Management of Chronic Kidney Disease in Adults in Primary and Secondary Care*. London, UK: National Institute for Health and Care Excellence; 2014.
11. Inker LA, Astor BC, Fox CH, et al. KDOQI US commentary on the 2012 KDIGO clinical practice guideline for the evaluation and management of CKD. *Am J Kidney Dis*. 2014;63(5):713-735.
12. Agrawal V, Ghosh AK, Barnes MA, McCullough PA. Awareness and knowledge of clinical practice guidelines for CKD among internal medicine residents: a national online survey. *Am J Kidney Dis*. 2008;52(6):1061-1069.
13. Smith SA. Estimation of glomerular filtration rate from the serum creatinine concentration. *Postgrad Med J*. 1988;64(749):204-208.
14. Boulware LE, Troll MU, Jaar BG, Myers DI, Powe NR. Identification and referral of patients with progressive CKD: a national study. *Am J Kidney Dis*. 2006;48(2):192-204.
15. Ozdemir S, Jafar TH, Choong LHL, Finkelstein EA. Family dynamics in a multi-ethnic Asian society: comparison of elderly CKD patients and their family caregivers experience with medical decision making for managing end stage kidney disease. *BMC Nephrol*. 2019;20(1):73.
16. Finkelstein EA, Ozdemir S, Malhotra C, Jafar TH, Choong HL, Suhardjono J. Identifying factors that influence physicians' recommendations for dialysis and conservative management in Indonesia. *Kidney Int Rep*. 2017;2(2):212-218.
17. Visser A, Dijkstra GJ, Huisman RM, Gansevoort RT, de Jong PE, Reijneveld SA. Differences between physicians in the likelihood of referral and acceptance of elderly patients for dialysis-influence of age and comorbidity. *Nephrol Dial Transplant*. 2007;22(11):3255-3261.
18. Hainmueller J, Hangartner D, Yamamoto T. Validating vignette and conjoint survey experiments against real-world behavior. *Proc Natl Acad Sci U S A*. 2015;112(8):2395-2400.
19. Peabody JW, Luck J, Glassman P, Dresselhaus TR, Lee M. Comparison of vignettes, standardized patients, and chart abstraction: a prospective validation study of 3 methods for measuring quality. *JAMA*. 2000;283(13):1715-1722.
20. Evans SC, Roberts MC, Keeley JW, et al. Vignette methodologies for studying clinicians' decision-making: validity, utility, and application in ICD-11 field studies. *Int J Clin Health Psychol*. 2015;15(2):160-170.
21. Pande SS, Pande SR, Parate VR, Nikam AP, Agrekar SH. Correlation between difficulty & discrimination indices of MCQs in formative exam in Physiology. *South-East Asian Journal of Medical Education*. 2013;7(1):45-50.
22. Rothman KJ, Greenland S, Walker AM. Concepts of interaction. *Am J Epidemiol*. 1980;112(4):467-470.
23. Knol MJ, van der Tweel I, Grobbee DE, Numans ME, Geerlings MI. Estimating interaction on an additive scale between continuous determinants in a logistic regression model. *Int J Epidemiol*. 2007;36(5):1111-1118.
24. Diaz-Quijano FA. A simple method for estimating relative risk using logistic regression. *BMC Med Res Methodol*. 2012;12:14.
25. Knol MJ, Le Cessie S, Algra A, Vandenbroucke JP, Groenwold RH. Overestimation of risk ratios by odds ratios in trials and cohort studies: alternatives to logistic regression. *CMAJ*. 2012;184(8):895-899.
26. Jhaveri KD, Sparks MA, Shah HH, et al. Why not nephrology? A survey of US internal medicine subspecialty fellows. *Am J Kidney Dis*. 2013;61(4):540-546.
27. National Kidney Foundation. K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Am J Kidney Dis*. 2002;39(2)(suppl 1):S1-S266.

28. Kagoma YK, Weir MA, Iansavichus AV, et al. Impact of estimated GFR reporting on patients, clinicians, and health-care systems: a systematic review. *Am J Kidney Dis*. 2011;57(4):592-601.
29. Wei L, MacDonald TM, Jennings C, Sheng X, Flynn RW, Murphy MJ. Estimated GFR reporting is associated with decreased nonsteroidal anti-inflammatory drug prescribing and increased renal function. *Kidney Int*. 2013;84(1):174-178.
30. Foote C, Clayton PA, Johnson DW, Jardine M, Snelling P, Cass A. Impact of estimated GFR reporting on late referral rates and practice patterns for end-stage kidney disease patients: a multilevel logistic regression analysis using the Australia and New Zealand Dialysis and Transplant Registry (ANZDATA). *Am J Kidney Dis*. 2014;64(3):359-366.
31. Feldman MD, Petersen AJ, Karliner LS, Tice JA. Who is responsible for evaluating the safety and effectiveness of medical devices? The role of independent technology assessment. *J Gen Intern Med*. 2008;23(suppl 1):57-63.
32. Hingwala J, Wojciechowski P, Hiebert B, et al. Risk-based triage for nephrology referrals using the kidney failure risk equation. *Can J Kidney Health Dis*. 2017;4:2054358117722782.
33. Pusic MV, Boutis K, Hatala R, Cook DA. Learning curves in health professions education. *Acad Med*. 2015;90(8):1034-1042.
34. Resnic FS, Wang TY, Arora N, et al. Quantifying the learning curve in the use of a novel vascular closure device: an analysis of the NCDR (National Cardiovascular Data Registry) CathPCI registry. *JACC Cardiovasc Interv*. 2012;5(1):82-89.
35. Maruthappu M, Gilbert BJ, El-Harasis MA, et al. The influence of volume and experience on individual surgical performance: a systematic review. *Ann Surg*. 2015;261(4):642-647.
36. Hemmelgarn BR, Manns BJ, Zhang J, et al. Association between multidisciplinary care and survival for elderly patients with chronic kidney disease. *J Am Soc Nephrol*. 2007;18(3):993-999.
37. Jones C, Roderick P, Harris S, Rogerson M. Decline in kidney function before and after nephrology referral and the effect on survival in moderate to advanced chronic kidney disease. *Nephrol Dial Transplant*. 2006;21(8):2133-2143.
38. Levin A, Adams E, Barrett BJ, et al. Canadians Seeking Solutions and Innovations to Overcome Chronic Kidney Disease (Can-SOLVE CKD): form and function. *Can J Kidney Health Dis*. 2018;5:2054358117749530.
39. Goldstein CE, Weijer C, Brehaut JC, et al. Ethical issues in pragmatic randomized controlled trials: a review of the recent literature identifies gaps in ethical argumentation. *BMC Med Ethics*. 2018;19(1):14.
40. Edwards SJ, Lilford RJ, Hewison J. The ethics of randomised controlled trials from the perspectives of patients, the public, and healthcare professionals. *BMJ*. 1998;317(7167):1209-1212.
41. Steiner PM, Atzmüller C, Su D. Designing valid and reliable vignette experiments for survey research: a case study on the fair gender income gap. *J Meth Measurement Soc Science*. 2016;7(2):52-94.
42. Murphy KR, Herr BM, Lockhart MC, Maguire E. Evaluating the performance of paper people. *J Appl Psychol*. 1986;71(4):654-661.
43. Langley GR, Tritchler DL, Llewellyn-Thomas HA, Till JE. Use of written cases to study factors associated with regional variations in referral rates. *J Clin Epidemiol*. 1991;44(4-5):391-402.
44. Lanza ML. A methodological approach to enhance external validity in simulation based research. *Issues Ment Health Nurs*. 1990;11(4):407-422.